Sex Differences and the Neural Correlates of Safety Learning:

Author: Allison R. Foilb

SEX DIFFERENCES AND THE NEURAL CORRELATES OF SAFETY LEARNING

ALLISON R. FOILB

A dissertation

submitted to the Faculty of

the department of Psychology

in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Boston College
Morrissey College of Arts and Sciences
Graduate School

January 2019

Title: Sex Differences And The Neural Correlates Of Safety Learning

By: Allison R. Foilb

Advisor: John P. Christianson, Ph.D.

Abstract: Accurate discrimination between safety and danger is necessary for survival, but is aberrant in individuals with post-traumatic stress disorder (PTSD). Despite its clinical relevance, very little is known about the cognitive and neural processes that underlie safety learning. Understanding how cues become safety signals is critical to understanding the impairments in fear modulation observed in individuals with PTSD. PTSD is more prevalent in women than men, and while research on sex differences in safety learning is limited, there is substantial evidence that males and females acquire and utilize safety signals differently. The aim of this dissertation is to describe behavioral sex differences in learning and recall of fear discrimination and explore the neural circuitry that allows this discrimination to occur.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 3

CHAPTER 4

CHAPTER 5

LIST OF ABBREVIATIONS

aIC	Anterior insular cortex
BLA	Basolateral amygdala
BNST	Bed nucleus of the stria terminalis
CeA	Central amygdala
IC	Insular cortex
IL	Infralimbic region of vmPFC
mIC	Medial insular cortex
NAcc	Nucleus accumbens
OFC	Orbitofrontal cortex
PFC	Prefrontal cortex
pIC	Posterior insular cortex
PL	Prelimbic region of vmPFC
VH	Ventral hippocampus
vlOFC	Ventrolateral orbitofrontal cortex
vmPFC	Ventral medial prefrontal cortex

ACKNOWLEDGEMENTS

I would like to begin by thanking my advisor, Dr. John Christianson, for making this research possible and helping me develop into the scientist and person I am today. Thank you for your instruction, guidance, and mentorship over the past five and a half years. I am truly grateful that I've had the opportunity to work with you and learn from you. I appreciate all of the time you have dedicated to my personal and professional growth.

Thank you to my dissertation committee—Drs. Michael McDannald, Maureen Ritchy, and Susan Sangha. I truly appreciate the time you have dedicated to supporting my work and helping me to improve as a writer, statistician, and scientist through your detailed and thoughtful feedback.

I would like to thank all of the past and current members of the Christianson lab for their ongoing scientific support, as well as their friendship. In particular, thank you to Nick, Morgan, Mary, and Anthony for your assistance with the work presented here. I would also like to mention and thank the undergraduates who assisted with this work – Mary, Veronica, Julia, Gab, and Emily – without you the scale of work I was able to accomplish would not have been possible.

Last but not least, thank you to my family and friends for your love and support throughout graduate school. Mom and Dad – thank you for always encouraging me and believing in me. Nick – thank you for enduring patience and support, and for your advice and feedback on science, writing and beyond. Your contribution to my success is immeasurable.

CHAPTER 1

Methods and brain mechanisms for learning and using safety signals

Portions of this chapter have been published in the following book chapter:

Foilb AR, Christianson JP (2018) Brain Mechanisms for Learning and Using Safety Signals. In: Neurobiology of Abnormal Emotion and Motivated Behaviors (Sangha S, Foti D, eds), pp 204–222.

1.1 Overview

Discriminating between safety and danger is elementary to the well being of species across the animal kingdom. Distinguishing predator from prey, poison from sustenance, and friend from foe are vital to survival. For humans, there are mundane discriminations made on a day-to-day basis—such as which traffic light means go and which means stop, or knowing under what circumstances being approached by a stranger could be dangerous compared to everyday interactions that we trust as safe. Most people make these safe versus dangerous discriminations with ease, and in many cases signals for safety outweigh signs of danger. Although proper safety judgments are important for our survival, the field of neuroscience does not have a very firm grasp on the neurobiological machinery that permits this function. This is in contrast to the detail known about the neural mechanisms of danger learning, which is reviewed later in this introduction. The disparity in knowledge between danger and safety is not for a lack of interest or novelty; Ivan Pavlov introduced the elementary consequence of safety learning—"conditioned inhibition" – nearly 100 years ago. Instead, the disparity exists because understanding danger systems is a prerequisite to understanding safety. Here I introduce learned safety signals and a neuroanatomical systems framework that organizes the known neural mechanisms and correlates of safety signals into a hypothetical model that is the basis of my dissertation work.

Before elaborating on the safety learning literature, it is useful to understand the significance of safety signals and their relevance to psychiatric conditions, namely post-traumatic stress disorder (PTSD). PTSD has a lifetime prevalence of 7.8% and represents a significant public health concern (Kessler et al., 1995; Kilpatrick et al., 2013). PTSD is often thought of as a disorder of over-fear conditioning, which can lead to hyper-vigilance and exaggerated physiological responses that are part of the clinical diagnosis for PTSD (Rauch et al., 2006). A hallmark feature of PTSD is the expression of fear or anxiety in environments where it is not appropriate or under conditions that would not typically elicit anxiety in a healthy individual (Rauch et al., 2000). This symptom has been conceptualized as a generalization of fear learned during trauma that becomes resistant to extinction (Rauch et al., 2006). The observations of generalized fear and fear extinction in individuals with PTSD have significant empirical support and trauma-induced changes in fear systems are hypothesized be central to the pathophysiology of PTSD. However, this hypothesis does not account for effects of trauma on an individual's capacity to use safety signals to regulate emotions. Indeed, difficulty utilizing learned safety signals has been observed in a number of clinical PTSD studies (for review, see Jovanovic et al., 2012). For example, in a visual task where a danger cue indicates to the participant that an aversive air puff is about to occur while a distinct safety cue indicates the absence of the air puff, healthy participants readily distinguish among the cues. This manifests as differential fear expression to the safe versus danger cues and *inhibition* of fear

when the fear and danger cues are presented together (Jovanovic et al., 2005). When individuals with PTSD are given the same training and testing, although they can report that the danger and safe cues are different, they are unable to reduce their fear when the safety signal is presented (Jovanovic et al., 2009, 2010). Thus, exposure to trauma and the development of PTSD compromises an individual's safety signal system, underscoring the importance and potential therapeutic impact of understanding how safety signals are learned and recalled in the brain.

Here the term safety signal refers to any cue that can, when presented in compound or in juxtaposition to fear evoking stimuli, reduce the behavioral or physiological expression of fear. Although there are innate safety cues that vary by species, for example a child's mother, this collection of work focuses on safety cues that are learned. Since the experiments of Pavlov on inhibitory learning, numerous experimental conditioning procedures have been set forth that result in safety signals. The following sections will summarize the behavioral paradigms that have been used to investigate the neural basis of safety signals, discuss the advantages of each with regard to understanding neural mechanisms, and review the information gained about the neural basis of safety signals. This work creates an incomplete picture of the neural circuitry involved in learning and using safety signals.

Associative learning processes permit an organism to remember environmental cues that predict danger or safety. One learns that a cue is

dangerous because it either occurred contemporaneous to or preceded an event that caused the subject harm. Subsequent presentations of the cue will elicit responses in the subject that prepare for impending danger. To study danger learning in a laboratory setting, Pavlovian fear conditioning procedures are used. In this procedure, a neutral stimulus (the conditioned stimulus, or CS) is paired with a mild electric shock (the unconditioned stimulus, or US). When the CS is later presented alone, it elicits fear, often observed as behavioral freezing in laboratory rats. Unlike danger learning, which is evident after a single CS-US pairing, safety learning occurs more gradually. Cues presented without consequence (i.e. no aversive US) under conditions when there is a non-zero probability that an aversive stimulus is imminent may become safety signals. Safety learning may constitute two processes; the first entails the discriminative learning that allows the subject to distinguish between the danger and safe cues. This fear discrimination is prerequisite to a special category of safety signals, which may be called *conditioned fear inhibitors*.

Based in part on the experimental and theoretical work of Pavlov (1927) and Konorski (1948), Rescorla (1969) outlined empirical requirements for establishing a conditioned inhibitor. Conditioned fear inhibitors must meet two requirements in addition to having a contrasting value to the conditioned excitatory cue—they must inhibit fear evoked during simultaneous presentation of a danger cue and show resistance to relearning as a danger cue. Presenting the excitatory cue and inhibitor in compound, a so-called “summation test”, is the

most direct method to test conditioned inhibition. In this sense, conditioned inhibition of fear occurs when a safety cue (often termed CS-, or B) indicating the absence of danger can reduce fear in the presence of a danger cue (often termed CS+, or A). In fact, the safety cue, B, must possess such a strong association contrasting the danger cue, A, that there is diminished fear learning when B is paired with the aversive US compared to a neutral and novel cue (Hammond, 1968). That is, once a cue is a conditioned inhibitor, pairing the inhibitor with an unconditioned excitor, such as shock, results in impaired excitation learning, or freezing, compared to a neutral cue. People (Jovanovic et al., 2005), monkeys (Winslow et al., 2008) and rats (Myers and Davis, 2004) are all able to discriminate between danger and safety in laboratory settings. The following section details the procedures used to train a subject to discriminate between safety and danger, and the behavioral tests used to assess conditioned fear inhibition by safety cues.

1.2 Models of Safety Learning

A number of methods are used to produce fear discrimination in research settings (See Table 1.1). The primary condition necessary for a stimulus to become a safety signal is that it occurs together with an excitatory fear stimulus, but is then followed by the absence of shock when shock is otherwise expected. In such circumstances, the value of a safety signal is apparent and allows for fear reduction because it predicts that the risk of harm is minimal in a situation when

an aversive event might occur. Safety signals can be established using a variety of approaches. In the conditioned inhibition procedure, named after the phenomenon itself, cue A is paired with a US on a given number of trials, while on other trials A is presented in compound with a second cue, B and the US is not presented. The presentation of cues A and B together forms a compound conditional stimulus. With training to this paradigm, subjects learn to fear A but not B, as B signals when A will not be followed by the US. The use of this procedure in the field has diminished greatly due to concerns that the novelty of two cues being presented in compound causes external inhibition, rather than conditioned inhibition (Myers and Davis, 2004).

Method	Description	Citations
Conditioned Inhibition	A+/AB- where A and B are presented in compound	Josselyn et al., 2005; Sangha et al., 2013
Feature Negative Discrimination	A+/BA- where B precedes A on non-reinforced trials	Falls & Davis, 1995; Falls et al., 1997; Heldt et al, 2002; Waddell et al., 2003; Heldt and Falls, 2006; Campeau et al., 1997
Differential Inhibition	A+/B-	Schiller et al., 2008; Genud-Gabai et al., 2013; Sangha et al., 2013; Likhtik et al., 2014; Stujenske et al., 2014
Differential Inhibition with Transfer Cue	AX+/BX- where X is presented in compound or in serial with A and B. X is a transfer cue which carries the expectation of danger on non-reinforced trials	Myers & Davis, 2004; Jovanovic et al., 2005, 2009, 2010; Winslow et al., 2008; Toufexis et al., 2007; Gutman et al., 2011; Foilb and Christianson, 2016; Chen et al., 2016; Foilb et al., 2016; Sarlitto et al., 2018
Explicitly Unpaired	B is temporally distant from reinforcement	Rogan et al., 2005; Pollak et al., 2008; Amano et al., 2010; Ostroff et al., 2010
Backwards Conditioning	B signals the end of the aversive reinforcement	Christianson et al., 2008; Christianson et al., 2011

Table 1.1 Laboratory procedures for safety learning. Schematics outline cue presentations, where red squares are danger cues (A), green squares are safety cues (B), blue squares are cues A and B presented in compound, and grey squares are transfer cues (X). Presentations of the aversive US are signified by lightning bolts (often electric shock or air puff in laboratory settings). Citations indicate relevant publications that utilize a particular method.

Feature-negative discrimination is a method similar to the conditioned inhibition procedure. Here, cue A is paired with a US on some trials, and on other trials B is a negative feature stimulus that precedes cue A. The serial presentation of B and A indicates the absence of shock, while A without the

negative feature stimulus continues to indicate shock. Notably, the feature stimulus on its own is not inhibitory, failing to meet Rescorla's definition of a conditioned inhibitor (1969). In these instances, the inhibitory strength of the feature stimulus is not through an inhibitory association with the US, but rather inhibits the excitatory value of the danger CS (Holland, 1984). In this case, the feature stimulus is "setting the occasion" for the CS to not indicate danger, rather than inhibiting the conditioned fear. While summation can still occur in occasion setting, it is know that occasion setting forms through separate neural mechanisms (for review, see Swartzentruber, 1995).

To test that a conditioned inhibitor is not an occasion setter, a transfer test can be performed. In this test, the negative feature stimulus is paired with a novel excitatory CS and inhibition of responding is measured. Occasion setters do not have inhibitory properties when presented with new excitatory cues, while conditioned inhibitors maintain their inhibitory properties (Holland, 1989). In some cases, the feature-negative procedure does lead conditioned inhibition, such that inhibitory properties of the negative-feature can transfer to a novel excitatory stimulus (Falls and Davis, 1997). Studies that successfully produce conditioned inhibitors often choose to show the inhibitory strength of the B cue through transfer tests or through a reduction in freezing when presented alone.

A popular method for producing cue discrimination and conditioned inhibition is differential inhibition. In this procedure, a discrete cue A is paired with an aversive US, while a discrete cue B is presented during the same conditioning

session but is never paired with the US. After repeated presentations of A+ and B-, the subject displays fear to A but not B, successfully discriminating between the two stimuli. Using this method, B often becomes a conditioned inhibitor. However in a summation test, where A and B are presented in compound, it is difficult to know if reduced fear is due to conditioned inhibition or due to external inhibition caused by the novelty of two simultaneous cues.

Conditional discrimination, often termed AX+/BX-, is a variation of differential inhibition paradigm that addresses the external inhibition problem. Originally used by Wagner et al. (1968) to study the associative strength acquired by X, it was later described my (Myers and Davis, 2004), as an ideal method to produce conditioned inhibition. To test external inhibition in this method, Myers and Davis varied the order and timing of cue presentations and found that simultaneous presentations of two cues in compound resulted in minimal external inhibition. Therefore, cue B becomes inhibitory through the same process as in the conditioned inhibition paradigm. Since cue X is compounded with the excitatory CS, and never presented on its own without the US, extinction of fear to cue X will not occur. X will therefore remain excitatory, and when B is paired with X, B becomes inhibitory as a result of being paired with excitatory cue X in the absence of the US. This method is successful in producing conditioned inhibition of fear while eliminating concerns of external inhibition that exist with the presentation of compound stimuli in a summation test. This procedure has

also been used to effectively produce fear discrimination and conditioned inhibition in healthy human participants (Jovanovic et al., 2005).

The work presented in this dissertation uses a serial variation of the AX+/BX- paradigm proposed by Myers and Davis (2004). In this procedure, the transfer cue X immediately precedes either the danger cue A or safety cue B during conditioning. This procedure allows for the same inhibitory learning to B from being paired with X as in the conditional discrimination procedure and creates significant safe/danger discrimination, as well as conditioned inhibition of fear, with no evidence of external inhibition. Our lab has displayed the efficacy of this procedure, through both a summation test and delayed learning as a danger cue (Foilb et al. 2016).

Additional methods for producing conditioned inhibition of fear are used in laboratory research. In an explicitly unpaired paradigm, a cue B is presented in the context where the aversive US is presented, but the B cue is temporally distal to the occurrence of the US. In this paradigm, the context and temporal cues serves as fear transfer stimuli since each predict a non-zero probability of US presentation. Because B and the US are never temporally paired, B has the potential to acquire inhibition. This method can also be used to compare to animals that undergo fear conditioning, where the tone and shock are paired. This method is similar to differential inhibition since the inhibitory value of the B cue may actually be in contrast to the excitatory association with context (Miller et al., 1991). Similarly, in backwards conditioning a CS is presented immediately

after the aversive US. In this procedure, the CS comes to signal the onset of a US-free period, giving it inhibitory associative strength.

Once learned, assessing the discriminative stimulus or fear inhibitor is achieved in one of several types of recall tests. A fear discrimination test simply entails presenting the subject with the danger and safe cues at separate times and observing fear evoked to each; safety signals will evoke significantly less fear than danger cues. In rats, fear is often assessed as freezing, a defensive behavior observed as complete immobility (Fanselow and Bolles, 1979; Fanselow, 1980, 1984), conditioned suppression of feeding (Estes, 1941; Hammond and Maser, 1970) or fear potentiated startle (Brown et al., 1951). In the work presented in this dissertation, freezing is used as behavioral measure of fear.

Two tests may be used to test for Rescorla's definition of a conditioned fear inhibitor. To test that a cue has gained inhibitory properties, Pavlov (1927) introduced the summation test. In summation tests, the putative conditioned fear inhibitor and a conditioned danger CS are presented in compound. If the cue is indeed a conditioned inhibitor, there will be significantly reduced fear to the AB compound cue compared to cue A alone. Hammond (1968) introduced the "retardation-of-acquisition test" in which the putative conditioned inhibitor is paired with an aversive US, essentially reversing the stimulus outcome expectancy present during initial conditioning. If a cue has become a conditioned

fear inhibitor, new danger learning will occur more slowly, presumably due to the preexisting inhibitory relationship between the safety cue and the US.

To summarize, there are several methods for cue presentation that will generate safe versus danger discriminations and safety cues capable of active fear inhibition. The following sections review the experiments conducted in rodents and primates to identify neural correlates of safety signals. In most cases, these studies have attempted to distinguish brain systems involved in the *acquisition* of safety signals (i.e. what brain regions are necessary for safety learning?) from brain systems involved in the *recall* and *use* of safety signals (i.e. what brain regions are needed for safety signals to inhibit fear?).

1.3 Mechanisms

1.3.1 Fear Circuitry

Because inhibition of a fear response, such as freezing, is central to the operational definition of a safety signal, understanding the brain mechanisms of danger learning and the expression of fear are prerequisite to understanding safety. Danger learning and fear expression have been extensively studied using Pavlovian fear conditioning procedures (McNally and Westbrook, 2006). Based on the work of Pavlov (1927), an innately neutral CS is paired with an aversive US. As a result of this pairing, a species-specific fear response is elicited by presentation of the CS. It is important to note that the majority of the methods for

achieving conditioned inhibition in Table 1.1 include CS-US pairings, so fear conditioning is inherent in almost all safety learning experiments.

Danger learning occurs primarily within the amygdala where neuroplasticity binds the CS and US in association as a result of their temporal contiguity. A subsystem of the amygdala termed the basolateral amygdala (BLA), consisting of the lateral, basolateral and basomedial nuclei, receives sensory inputs, including nociceptive information, from diverse brain areas, including the thalamus, neocortex, olfactory cortex, and hippocampus (Turner and Zimmer, 1984; LeDoux et al., 1990; Stefanacci et al., 1992; Mascagni et al., 1993; Romanski and LeDoux, 1993; LeDoux, 1996; McDonald, 1998; Kim and Jung, 2006). This makes the BLA a likely site of convergence for information about the CS and US (LeDoux, 2000). Indeed, when the CS and US are coincident, synaptic plasticity occurs in the BLA such that subsequent presentations of the CS alone evoke stronger BLA activation than would an unconditioned CS (Quirk et al., 1995; Rogan et al., 1997). Accordingly, manipulations that prevent BLA excitability or plasticity interfere with the learning and later expression of conditioned fear (Maren et al., 1996; Cousens and Otto, 1998; Lalumiere, 2014). Excitation within the BLA begins a cascade of circuit activation via projections to the central amygdala (CeA) and bed nucleus of the stria terminalis (BNST), which in turn project to the hypothalamus and brainstem areas which are the proximate mediators of specific fear responses (LeDoux et al., 1988; Swanson and Petrovich, 1998; Maren, 2001) including freezing, autonomic arousal, hormone

release, analgesia, and startle (LeDoux et al., 1988; Van de Kar et al., 1991; Davis, 1992; Kapp, 1992). More specifically, the projection from BLA to CeA mediates fear responses to cues of short duration; whereas prolonged fear responses are mediated by BLA to BNST projections (Davis et al., 2010).

A fundamental assumption taken when considering the mechanisms of safety signals is that they inhibit fear responses by modulation of this fear circuitry. Simply put, a safety signal might operate by blunting activity or flow of excitation through the fear circuit at the sensory (i.e. thalamic), associative/learning (i.e, BLA) or output phases (i.e, CeA, BNST or brainstem nuclei).

1.3.2 Safety Signals in Basolateral Amygdala

As noted, the BLA is the site of neuroplasticity for fear learning, and it is necessary for the learning and expression of conditioned fear. It is then reasonable to predict that safety signals might also utilize the BLA for both learning and recall, where a safety signal would be expected to reduce BLA activity compared to danger signals. Many studies have in fact found evidence that safety signals impact responding in the BLA. The various approaches that have been used to study the role of BLA in safety learning further substantiate the role of BLA in this process.

Using monkeys and a differential inhibition procedure, Genud-Gabai et al. (2013) paired a tone or visual cue with an aversive air puff near the eye, while a

different tone and unique visual cue served as B cues, indicating the absence of the US. Eye blink was used as a behavioral measure of fear and monkeys significantly discriminated between the A and B cues. Learning to associate B cues with the aversive US was also delayed. Microelectrodes placed above the amygdala for neuronal single unit recording during acquisition of this fear discrimination found amygdala neurons firing in the presence of both A and B cues. In trials where more neurons fired to A than B, there was increased generalization of fear. This study indicates that amygdala encoding of discriminative CSs is important to fear discrimination and safety learning.

Sangha et al. (2013) found similar results when looking at safety encoding in the amygdala of rats. A combination of conditioned inhibition procedure and differential inhibition methods were used during sessions of *in vivo* single unit recording, where A trials were paired with shock and the AB compound without shock, as well as some trials where B was presented alone in the absence of shock. Retardation of fear acquisition to B was also displayed to verify its value as a conditioned inhibitor. Over a third of neurons sampled altered firing selectively to either fear or safety stimuli. About a quarter of these neurons selectively altered firing during safe signals—the AB compound or B alone—particularly during later conditioning sessions. Several neurons also showed significant inhibition to safe signals.

Ostroff et al. (2010) looked at the effect of safety learning on amygdala neurons by comparing spine morphology of lateral amygdala neurons in animals

that underwent fear conditioning to animals that received an explicitly unpaired procedure to produce conditioned inhibition, which passed both summation and retardation tests. They found increased synapse size with fear conditioning and decreased synapse size with safety learning. Further investigation of these neurons found that while amygdala synapses strengthen after fear conditioning, they weaken when conditioned inhibition is established (Ostroff et al., 2012). These bidirectional changes based on fear associations indicates that both fear and safety learning alter synaptic morphology in the amygdala.

The above investigations of amygdala function in safety used correlational methods. This is perhaps the only reasonable approach to investigating the amygdala because any manipulation that would impair amygdala function, and therefore allow a mechanistic interpretation, would necessarily impair either danger learning or recall and preclude an observation of fear inhibition by a safety signal. Nevertheless, Kazama et al. (2012) performed neonatal amygdala lesions in rhesus monkeys that were then trained as adults on the differential inhibition paradigm adapted for monkeys, with fear potentiated startle as a measure of fear (Winslow et al., 2008). Neonatal amygdala lesions delayed the acquisition of learned fear, but did not impair discrimination between safety and danger, or summation. These results indicate that conditioned inhibition in the fear potentiated startle paradigm may develop independent of amygdala activation. The current literature paints an incomplete picture of amygdala function in safety learning but it is clear that neuronal firing within the BLA

differentiates between safety and danger. Whether or not these units, or plasticity within the BLA, are required for safety learning or later fear inhibition has not yet been tested with a direct, mechanistic approach.

1.3.3 Safety Signals and Fear Expression Circuits

The principle outputs of the BLA that initiate and maintain fear responses are the CeA and the BNST. The CeA receives sensory and visceral information from the BLA, and projects to the hypothalamus and brainstem areas responsible for the fear response (LeDoux et al., 1988; Swanson and Petrovich, 1998; Maren, 2001). Falls and Davis (1995) made lesions to the CeA after extensive training using the feature-negative discrimination procedure. Since lesions of the CeA block the expression of fear-potentiated startle to A, additional A and shock training was conducted until fear returned. Without receiving additional training with the B cue, animals with CeA lesions were able to inhibit fear-potentiated startle to the AB compound, indicating that CeA is not critical for the expression of conditioned of fear (Falls and Davis, 1995).

Campeau et al. (1997) used immediate early gene protein product Fos to quantify neuronal activation following feature negative conditioning. Fear was measured via potentiated startle, where startle was reduced during presentation of AB and B compared to A. Presentation of the AB and B cues led to Fos expression in the dorsal caudate nucleus of the striatum and BNST. Providing a backwards CS also leads to differential activation of the BNST, without effect on

the CeA (Christianson et al., 2010). The role of BNST in this mechanism may indicate why others have found no effect of CeA lesions in the expression of safety learning (Falls and Davis, 1995; Kazama et al., 2012). BNST efferents are very similar to those of the CeA and so this region is involved in the sustained expression of fear (for review, see Walker and Davis, 2008). BNST may provide a unique or redundant mechanism for fear expression in these paradigms that is specifically regulated by learned safety signals. A better understanding of how safety signals affect the different nuclei and neuronal subgroups within the BNST should be a rich area for future research.

1.3.4 Safety Signals and Sensory Systems

As safety signals must be encoded into the central nervous system through sensory systems, it is easy to assume that interfering with the subject's sensory capacity hear, see, smell or touch the safety signal would impede safety learning and recall. Falls and colleagues employed a feature-negative paradigm with a combination of visual and auditory cues as the CSs to test the role of sensory systems. Somewhat surprisingly, neither lesions to the auditory thalamus (Heldt and Falls, 1998) or perirhinal cortex (Falls et al., 1997) affected AB summation tests. However, Waddell, Heldt & Falls (2003) later found that lesions of the superior colliculus, a brainstem center for visual processing, prevented inhibition in a summation test. Interestingly, the inhibitory stimulus in this experiment was auditory cue. A complementary study found that lesions to the

inferior colliculus prevented the expression of summation to an auditory cue suggesting damage to fibers of passage could account for the interference caused by superior colliculus lesions (Heldt and Falls, 2003). While interruption of sensory relay is a parsimonious account of these studies, it is not yet possible to rule out a role for elementary sensory structures in safety signal learning or recall.

Although the results of the studies of Falls and colleagues suggest the thalamus may not be a *necessary* component for safety signal recall, a set of experiments by Rogan and colleagues (2005) suggests that thalamic inputs to the BLA are differentially altered by either danger or safety cues. Mice were trained in an explicitly unpaired paradigm or a conventional CS-US fear conditioning paradigm. Cue evoked potentials in the BLA were recorded in awake, behaving mice before and after unpaired or paired conditioning. Paired conditioning appeared to potentiate auditory evoked responses, suggesting a strengthening of the representation of tone in the thalamic-BLA tract, which had been reported earlier (Rogan et al., 1997). Conversely, when the tone indicated safety, the response in the BLA from auditory inputs was diminished, suggesting depotentiation at auditory-BLA synapses. While the experimenter's interpret their finding as a change in the thalamic input to the BLA, it is not yet known at this time whether the change in auditory evoked potential reflects reductions in specific thalamic or cortical sensory inputs to the BLA. This result indicates that safety signals may reduce the ability of a sensory cue to excite the BLA.

When a cue is temporally correlated with a sense of pain, then that cue becomes a danger signal, whereas if it is not paired with pain, or is paired with a sense of relief, it may become a safety signal. Therefore, computing whether a signal is safe or dangerous requires the detection of coincident sensory states across multiple stimulus modalities including the somatic, or interoceptive sensory system. Our lab hypothesized that insular cortex may play a role in conditioned inhibition because of its known access to somatosensory information, role in convergent response to multisensory stimuli (Rodgers et al., 2008) and bidirectional amygdala connectivity (Shi and Cassell, 1998a, b). In the first experiments, Christianson et al. (2008, 2011) reported that providing a backwards CS during a series of unpredictable shocks served as a safety signal and prevented the development of learned helplessness behaviors that typically occur when the safety signal is not present. Lesions and temporary pharmacological inactivation of a region of posterior insular cortex, identified for its capacity to integrate stimuli from multiple modalities (Benison et al., 2007; Rodgers et al., 2008), completely eliminated the stress protective effects of the safety signal (Christianson et al., 2008, 2011).

Recently, we reported a series of experiments using the serial differential inhibition procedure to determine if the insular cortex contributed to learning about safety signals or later summation (Foilb et al., 2016). In these studies (elaborated on in Chapter 2), blockade of NMDA receptors in the posterior insula prevented acquisition of inhibition of fear by the safety signal on later AB

summation tests. Interestingly, when rats were trained drug free, inhibition of insular cortex before a summation test reduced fear expression yet did not influence fear discrimination or conditioned inhibition of fear. Thus, the insular cortex seems to be important for learning about safety signals, but not for their recall. In control experiments, we found that the contribution of insula to safety learning could not be reduced to a simple function in fear discrimination. These results suggest that the insular cortex may contribute to safety learning not by encoding the safety signal, but by maintaining a stable representation of the danger signal. Together, studies on insular cortex clearly implicate its role in safety signal learning.

1.3.5 Fear Modulatory Circuits: Prefrontal Cortex, Hippocampus, & Striatum

In addition to sensory inputs, the amygdala is interconnected with structures that are known to influence the learning and expression of fear including the prefrontal cortex, dorsal and ventral hippocampus, and regions of striatum. These structures are critical for executive function, episodic memory, and reward seeking behaviors, respectively (for reviews see Phelps, 2004; Kesner and Churchwell, 2011; Hart et al., 2014). The ventral medial prefrontal cortex (vmPFC) plays a critical role in the extinction of fear (for review, see Milad et al., 2014), which has often thought to be closely related to the inhibition of fear by a safety signal. Further, bilateral lesions of mPFC in dogs disrupted conditioned inhibition of appetitive conditioning (Konorski, 1967). To investigate

the role of the vmPFC in conditioned inhibition of fear, Gewirtz, Falls and Davis (1997) used a feature-negative discrimination procedure to train and test animals before performing lesions to the vmPFC and then training and testing again. Quite surprisingly, vmPFC lesions appeared to have no effect, as inhibition of fear was evident in both lesioned and sham-lesioned rats on summation trials. Although learning and fear inhibition may have been achieved after vmPFC lesion by redundant or compensatory neural circuits, more recent investigations of vmPFC function are consistent with these initial results. Temporary inactivation of the vmPFC had no effect with a backwards conditioned CS (Christianson et al., 2008). Conversely, Vieira et al. (2015) used mutant mice to delete an essential subunit of N-methyl-D-aspartate (NMDA) receptors, known to be necessary for synaptic plasticity (Morris, 2013), in excitatory neurons in the mPFC and observed impaired discrimination between auditory A and B cues, with increased fear to the CS- in knockout mice compared to controls. These results indicate that plasticity in mPFC may be necessary for acquisition of fear discrimination.

Sangha et al. (2014) dissected the contributions of prefrontal cortex subregions—prelimibic (PL) and infralimbic (IL)—using a combination of the differential and conditioned inhibition procedures. Inactivation of PL led to a reduction of freezing to the danger cue A, but did not alter freezing to B or AB cues compared to vehicle animals. This result fit with the existing knowledge that PL mediates the heightened fear response in the presence of a danger cue

(Sotres-Bayon and Quirk, 2010). Inactivation of IL before recall testing resulted in reduced freezing to A, which abolished discrimination between the A and AB cues. These results are especially interesting in light of compelling support for a role of the IL in fear reduction after extinction (Sierra-Mercado et al., 2011). The results of these mechanistic experiments demonstrate that the vmPFC is not necessary for learning about safety *per se*, but vmPFC may play a role in distinguishing between what is dangerous and what is not.

Likhtik et al. (2014) examined the relationship of vmPFC and BLA synchrony during a differential inhibition task in mice. Recordings of local field potentials in the vmPFC and BLA revealed differential responses to learned danger and safety cues, with larger amplitude responses occurring in vmPFC, primarily the PL, to the danger cue. This apparent differentiation was evident in the strength of synchrony between vmPFC and BLA as it correlated to the difference in behavioral freezing to either the danger or safe cue, such that mice with very good behavioral discrimination exhibited the strongest connectivity between PFC and BLA. Consistent with Sangha's results, the role of the vmPFC appears to be in determining what is dangerous rather than what is safe.

Our lab also looked at the role of ventrolateral orbitofrontal cortex (vlOFC) in fear discrimination using a serial differential inhibition procedure, which is further described in Chapter 2 (Sarlitto et al., 2018). vlOFC has been implicated in value-based decision making (Sul et al., 2010), as well as in response selection (for review, see Young and Shapiro, 2011) and switching between

cognitive tasks (Wilson et al., 2014). Based on these functions we hypothesized that vlOFC would be recruited during fear discrimination recall to facilitate changes in behavioral freezing to safety and danger cues. Temporary inactivation of the vlOFC before a discrimination recall test impaired discrimination, resulting in greater fear to the safety cue B than vehicles. Inactivation of vlOFC during the serial differential inhibition conditioning procedure did not impact discrimination during acquisition or later recall. While future work is required, there is evidence that the vmPFC and vlOFC, contribute to different aspects of recall of both danger and safety signals, perhaps a consequence of a more general function in response selection.

Hippocampal regions have also been implicated in fear discrimination behavior. Pre-training lesions to the hippocampus did not impact discrimination performance in a feature-negative task, but post-training lesions impaired safety recall, such that there was no reduction of fear when B was presented in compound with A (Heldt et al., 2002). Heldt and colleagues did note that despite the deficits of post-training lesions, animals could successfully be retrained to make this discrimination. Our lab followed up on these results with a focus on ventral hippocampus, work that is further described in Chapter 2. Temporary inhibition of the ventral hippocampus prior to conditioning, prevented danger learning as subsequent presentations of A and B cues evoked little fear. Rats were later retrained and the ventral hippocampus was inactivated prior to a discrimination recall test, but no effect of ventral hippocampus inactivation was

apparent (Chen et al., 2016). Although contradictory to the results of Heldt et al. (2002), these results support the existing literature implicating a role for ventral hippocampus in fear acquisition (see Anagnostaras et al., 2001 for review), as well as to the discrimination of fear contexts (Orsini et al., 2011). While the results of this study indicate that ventral hippocampus may be part of the fear circuit, it doesn't appear to directly encode excitatory and inhibitory associations of discrete CSs.

Like Heldt et al. (2002), Pollak and colleagues (2008) found evidence for a role of hippocampus in safety learning. Comparing animals that underwent an explicitly unpaired procedure to those that experienced fear conditioning, they found that animals in the safety learning condition had increased hippocampal newborn cell survival, with no changes in neurogenesis, compared to fear conditioned animals. Ablation of hippocampal neurogenesis by X-irradiation delayed safety learning and prevented correlated "antidepressant" behavioral effects. With this study it is worth considering whether the explicitly unpaired paradigm recruited the hippocampus because safety is learned as a consequence of the temporally distance between the aversive US and the safety cue which could be mediated by the hippocampus, whereas differential conditioning or feature negative procedures, that appear to be hippocampal-independent, the temporal distance between cues that predict danger (transfer cues) and safety cues is small, often overlapping.

Josselyn and colleagues (2005) hypothesized that the nucleus accumbens

(NAcc) may be necessary for the fear modulating effects of a safety signal. The NAcc plays a role in modulating the motivational responding in appetitive conditioning (for review, see Castro et al., 2015) and is situated to perform a similar task in fear and safety learning, as it receives information from many of the neural structures involved in conditioned fear (McDonald, 1991). With this information, they tested the role of the NAcc in the conditioned inhibition procedure, pairing A with shock in phase one, and an A and B compound with no shock in phase 2. Using three independent manipulations of NAcc activity: lesion, AMPA receptor blockade, or amphetamine injection, no role for NAcc was found to alter the fear potentiated startle response, or conditioned inhibition of startle during AB summation trials.

Also focused on striatum, Rogan et al. (2005) recorded tone-evoked synaptic responses in the caudate putamen of mice in an explicitly unpaired paradigm. In contrast to auditory evoked responses to the safety signal in the amygdala, which decrease after conditioning, in the caudate putamen, tone-responses were enhanced with safety conditioning and weakened with fear conditioning. This was interpreted as plasticity associated with approach and reward, but the necessity of caudate putamen, or any other striatal region outside of the NAcc requires further mechanistic inquiry.

1.3.6 Neurotransmitter Systems

Serotonin plays a role in conditioned inhibition of appetitive learning (Lister et al., 1996) and conditioned analgesia (Watkins et al., 1998). In each case, destruction of serotonergic neurons impaired the effect of a conditioned inhibitor. Regarding fear discrimination, Berg et al. (2014) reported impairment in differential learning to a partially reinforced safety signal after lesions to the serotonergic dorsal raphe nucleus (DRN). This work suggests a role of the DRN, and likely serotonin, in using prediction errors to update associations between CS and US in discrimination learning. This learning mechanism could be important for safety learning in general. There is substantial literature implicating serotonin (5-HT) in the modulation of fear, with the general consensus that 5-HT release, and action at 5-HT_{2C} receptors in the amygdala enhances the expression of fear (Martin et al., 2002; Campbell and Merchant, 2003; Greenwood et al., 2008; Christianson et al., 2010; Baratta et al., 2016). We hypothesized that safety learning could be enhanced by reducing the fear enhancing effects of 5-HT_{2C} with specific receptor antagonists. Using the serial differential inhibition procedure, we blocked receptor subtype 5-HT_{2C} prior to conditioning, which resulted in improved fear discrimination, with more pronounced conditioned inhibition in summation tests (Foilb and Christianson, 2016). Given the vast number of neurotransmitter systems implemented in the modulation of fear and anxiety, future investigations of drugs that could reduce fear, may make for useful therapeutics to augment safety learning.

Dopamine is primarily thought of for its role in reward circuitry and forming cue-reward associations, which may be closely related to safety learning (Berridge and Robinson, 1998). It is logical that the presentation of a safety cue is a rewarding event when there is a non-zero probability of danger. Dopamine is also known to play a role in fear learning (for review, see Lee et al., 2016). In the amygdala, dopamine receptor activation is necessary for fear memory, dopamine in the BLA increases fear expression and D1 receptors of the BLA are required for inhibition of fear in extinction paradigms (Hikind and Maroun, 2008; de Oliveira et al., 2011; Lee et al., 2016). Receptors for dopamine, D1 and D2, have also been found important for modulating activity in uncertainty paradigms (Larkin et al., 2016). In safety learning, systemic administration of D1 agonist or antagonist impairs fear inhibition to a safety cue, with similar results observed when the D1 agonist and antagonist were administered intra-BLA (Ng et al., 2018). These results implicate dopamine as another potential target for treatment in disorders of appropriate fear modulation to a safety cue.

Sex differences have also been found in safety learning, indicating a potential role for sex-related hormones in facilitating fear discrimination. Using AX+/BX- discrimination conditioning, Toufexis et al. (2007) looked at the role of estrogen on fear and discrimination learning in both male and female gonadectomized rats. While the addition of estrogen did not alter discrimination in males, in females, estrogen prevented inhibition of fear to the AB compound. Implantation slow release capsules for estrogen receptor agonists disrupted fear

discrimination learning in both males and females, in a manner such that different types of estrogen receptors may alter a different aspect of discrimination learning. Animals given the agonist for estrogen receptor α displayed increased fear to all cues compared to animals given estrogen receptor β agonist. This also fits with existing data on estrogen receptors, where estrogen receptor α has been implicated in increased fear and anxiety and estrogen receptor β has been shown to be anxiolytic (Morgan and Pfaff, 2001; Walf et al., 2004; Walf and Frye, 2005). Further investigation of the role of sex-related hormones in safety learning may help explain why more women than men are diagnosed with PTSD, and how to address the inability to properly inhibit fear responses that is seen in females with PTSD (Jovanovic and Norrholm, 2011; McLean et al., 2011; Lebron-Milad and Milad, 2012).

1.4 Studies of Safety Signals in Humans

As noted in the review of safety learning methods, the conditional discrimination protocol allows for comparative studies of discrimination and safety learning between rodents and humans. Jovanovic and colleagues (2005) used this paradigm with healthy human participants where compounds of colored lights are presented either with an aversive air blast to larynx (AX+) or without an air blast (BX-). Tests for conditioned inhibition consisted of presentations of AX, BX, and AB in summation test, as well as presentations of A in compound with a novel stimulus. In healthy human participants, as in rodents, presentations of AB

led to significantly less potentiated startle than presentations of AX or A and a novel stimulus in compound. This indicates that inhibition of fear to the compound cue was due to its inhibitory association rather than external inhibition. This study provided clear evidence that healthy humans can modulate fear responding in the presence of safety signals using the same methods that are used to study animal models of fear modulation, which can facilitate translation of information between investigations from different species.

Jovanovic and colleagues went on to use this paradigm in civilian and combat PTSD patients (Jovanovic et al., 2009, 2010). In these studies, fear potentiation responses showed that the healthy controls and low-symptom PTSD participants successfully discriminated between safe and danger trials, but high-symptom PTSD subjects did not. The high-symptom PTSD group showed signs of discrimination, but the difference in startle response between AX and BX trials was minimal overall and not statistically significant. Similarly, controls and low-symptom PTSD subject displayed significantly less fear to AB than AX, while the high-symptom PTSD subjects showed similar levels of fear to AB and AX. In addition to recording fear-potentiated startle, participants were also asked to report their expectation that air blast, the danger CS, would occur on any particular trial. In the last block of conditioning, all groups showed successful discrimination based on their reported expectancy that the air blast would occur, indicating that high-symptom PTSD patients were able to learn about safety signals, but are unable to inhibit their physiological fear response. Further work

has shown that the safety signal deficiencies seen in individuals with PTSD may appear as early as 30 days after trauma and as late as 10 years after trauma exposure, indicating that it is a persistent biomarker of psychopathology (Jovanovic et al., 2013).

Although fear extinction has received considerable attention in fMRI studies, safety signals are less common in human neuroimaging. Schiller et al. (2008) studied the neural correlates of fear reversal with fMRI, and in the process, learned a little bit about the regions activated in the presence of safety and danger signals. The procedure used images of two different angry faces as cues, where face A was paired with a mild wrist shock on one third of trials in order to delay acquisition, and face B was never paired with shock. During both early and late acquisition of safe and danger associations, the B cue elicited stronger vmPFC responses than A. Unsurprisingly, amygdala activation was greater during presentations of the A cue than during presentations of B. Similarly, striatal responses were higher to the A cue compared to B during acquisition, as was activation in the caudate putamen, thalamus, midbrain, dorsal anterior cingulate cortex, superior frontal gyrus and insular cortex. Since the goal of this research was not focused on safe/danger discrimination, but rather when the meaning of stimuli change from safe to unsafe, they did not test for recall of the safe and danger stimuli to look at different brain regions implicated in the recall of fear discrimination.

Insular cortex is a region implicated in the pathophysiology of PTSD (Chen

et al., 2006) and, of the brain regions identified in studies of safety signals, it is the only one where there are consistent findings from both rodent and human studies. The fMRI results from (Schiller et al., 2008) showed significantly greater responding to A than B in the insular cortex during discrimination acquisition which is consistent with the effects of insula inactivation on fear expression in our hands (Foilb et al., 2016). Individuals with PTSD also show altered activity in insular cortex (Strigo et al., 2010) and, although preliminary, work from Gutman et al. (2010) correlated insular cortex volume with fear inhibition by safety signals. With female participants who met criteria for PTSD, they used the differential inhibition protocol combined with structural MRI to find that individuals with a higher startle response to the safety signal (that is, poor inhibitors) had smaller insula volumes than those that attenuated their fear response in the presence of the safety signal. Future studies must evaluate insular cortex activity using fMRI and connectivity methods to better elucidate the circuits that are required to learn about safety and recall of these cues to inhibit fear responses.

1.5 Sex Differences in Safety Learning

Much of the research on fear discrimination has focused on males, yet females are more likely to be diagnosed with PTSD than males (Kessler et al., 1995; Kilpatrick et al., 2013). The limited research on fear discrimination in females indicates that they may differ from males in how they learn about safety and generalize between cues indicating safety and danger. Research in humans

has found that fear discrimination is more sensitive to trauma history in females compared with males. Studying trauma-exposed, 8 to 13 year-old children, Gamwell and colleagues (2015) used a differential inhibition paradigm with an air blast US and fear-potentiated startle as a measure of fear, as well as asking children to respond whether or not the cue indicated an air blast. Despite higher rates trauma-exposure in boys, girls showed less discrimination between presentations of A and B. Interestingly, girls also habituated to A+ trials earlier than boys, showing little startle or skin conductance response by the end of conditioning. Lonsdorf et al. (2015) compared adult men and women in a context-dependent fear discrimination task and explored the role of menstrual cycle and hormonal contraceptives. For both the danger and safety cues, women reported higher levels of fear and US expectancy than men, while men had higher skin conductance responses to the danger cue compared to women. Overall, women displayed less discrimination between danger and safety cues compared to men, although this sex difference was mostly influenced by women on hormonal birth control, further indicating that sex hormones may play a critical role in sex differences in discrimination (Toufexis et al., 2007; Lonsdorf et al., 2015). However, menstrual cycle phase of free-cycling women did not impact fear discrimination. For skin conductance responses, free-cycling women showed similar discrimination to men, while women on hormonal birth control displayed significantly less fear discrimination. These results are particularly interesting in relation to how sex differences are often studied in rodents, with artificial levels of

sex-related hormones, namely estrogen, rather than the naturally fluctuating levels of these hormones through the estrous cycle.

Translational research regarding sex differences in rodent fear discrimination is in its infancy, but has yielded mixed results. Day et al. (2016) tested cued discrimination in rats and found greater discrimination in females compared with males. However, after repeated testing, females displayed generalization of fear, while males did not. Females also failed to display resistance to re-learning the safety signal as a danger cue, which is one of the two classic tests for conditioned inhibition, as defined by Rescorla (1969). In mice, females showed greater context fear generalization than males. Using Fos as a marker of neural activation, generalization in females corresponded with decreased activation in hippocampus and increased activation in amygdala compare to males (Keiser et al., 2017). While differing results in mice and rats is surprising, this may be due to differences in paradigms rather than differences in species, since Day et al. used cued discrimination and Keiser et al. focused on context discrimination. There is additional evidence that females overgeneralize between contexts, which may explain the results found by Keiser et al. (Lynch et al., 2013; Anderson and Petrovich, 2017; Keiser et al., 2017).

Describing the behavioral sex differences in fear discrimination and conditioned inhibition of fear is a main focus of this disseratation, and the results of these experiments are presented in Chapter 3. Biological sex is a significant factor in the expression of fear-based psychoses and a better understanding

basic behavioral sex differences in fear discrimination is needed to more fully realize the impact of sex for an individual's health (Shansky, 2015; Shansky and Woolley, 2016). Further investigation into sex differences in safety learning may explain why more women than men are diagnosed with PTSD, and address the inability of individuals with PTSD to properly inhibit fear responses - particularly if the mechanisms in females differ from those in males (Jovanovic and Norrholm, 2011; McLean et al., 2011; Lebron-Milad et al., 2012). This dissertation aims to thoroughly examine sex differences in safety learning and the underlying neural mechanisms; steps critical to progress in the treatment of PTSD and other anxiety disorders with impairments in fear modulation.

1.6 Aims of Dissertation

The overarching goal of this dissertation is to uncover the neural mechanisms that underlie the learning and recall of safety signals, as well as to address the important issue of sex differences in discrimination learning and recall, behaviorally. Existing data on potential sex differences in fear discrimination are inconclusive and neural correlates of behavioral sex differences are limited (Lonsdorf et al., 2015; Day et al., 2016; Keiser et al., 2017). Observing, first, how males and females learn, recall and utilize safety signals will provide insight on potentially unknown sex differences in behavioral learning and memory. Further exploring the brain regions that underlie discrimination learning in males and females will allow for a better understanding

of the neural circuits necessary for safety learning, as well as regions that may inhibit appropriate fear modulation.

Chapter 2 will review my previous work on neural mechanisms that underlie safety learning, using the same procedural methods that will be used throughout the remainder of my work presented here. In Chapter 2, we investigate the roles of ventral hippocampus and vlOFC in the acquisition and recall of safe/danger discriminations (Chen et al., 2016; Sarlitto et al., 2018). We found inhibition of ventral hippocampus reduced overall fear expression, but did not directly impact fear discrimination, while vlOFC inhibition specifically altered fear to the safety cue. Chapter 2 also includes research focused on the role of insular cortex and a paradigm for conditioned inhibition of fear. As previously described, insula is a region involved in sensory integration and densely connected to amygdala (Shi and Cassell, 1998a, b; Rodgers et al., 2008). We hypothesized that this structure may play a role in the integration of cues necessary for modulation of fear in the presence of safety cues. We found posterior insular cortex is specifically necessary for acquisition of conditioned inhibition of fear, but not for the acquisition of fear discrimination or the recall of either fear discrimination or conditioned inhibition as measured by a summation test.

While my previous work elaborated on the limited known neural correlates of safety learning, only male rats were used in these studies. Due to the prevalence of PTSD in women, the existing data indicating that males and

females may differ in fear discrimination abilities and the effect of estrogen on fear discrimination, I set out to explore potential sex differences in the learning and utilization of safety cues (Kessler et al., 1995; Toufexis et al., 2007; Kilpatrick et al., 2013; Day et al., 2016). Chapter 3 uses a large sample size to assess potential sex differences in learning to discriminate between danger and safety cues. Looking at conditioning by trial blocks to allow for precise analysis of how the sexes, on average, display differential responding to the A and B cues on over the course of the conditioning session, as well as average fear to the safe and danger cues during conditioning. A subset of animals received fear discrimination recall testing. Here I compare the average expression of fear to the safe and danger cues, as well as look at changes across presentations of the cues throughout the duration of the test. Another cohort of animals received repeated conditioning and summation tests, as described in Chapter 2. Surprisingly, we found that females reduced freezing to safe cue presentations earlier than males, and maintained lower levels of freezing to the safety cue throughout conditioning. Females also displayed reduced fear to the safety cue at the start of the discrimination recall test the subsequent morning. Conversely, in summation tests of conditioned inhibition, males and females did not differ in patterns of freezing to compound cue and there was no significant difference between sexes.

Since the sex difference appeared initially and most drastically during acquisition of fear discrimination, I was particularly interested what brain regions

may underlie safety learning and contribute to this behavioral sex difference. Chapter 4 assesses the neural structures that underlie acquisition of fear discrimination, using an immunohistochemistry (IHC) for immediate early gene product Fos as a neural marker of activation. Sex differences in neural activation of fear discrimination learning could indicate potential pathways for intervention in individuals with impairments in utilization of safety cues. This study used three groups of animals so that structures involved in fear discrimination could be distinguished from those involved in the sensory processes of cue discrimination and those that mediate fear learning when a safety signal is not present. Based on existing information about the neural structures involved in fear learning and fear discrimination described here, I composed a hypothetical neural circuitry (Figure 1; Foilb and Christianson, 2018) that may underlie discrimination learning and pinpoints the brain regions of interest that are the focus of Fos investigation. It is important to look at regions known to be involved in fear learning and expression, such as BLA, BNST and CeA, particularly since there is also evidence that these regions may play critical roles in safety. Regions known to be involved in the modulation of fear, including PL and IL, vlOFC and insular cortex, are also key structures of investigation. These brain regions are of particular interest because their known involvement in fear and safety circuitry.

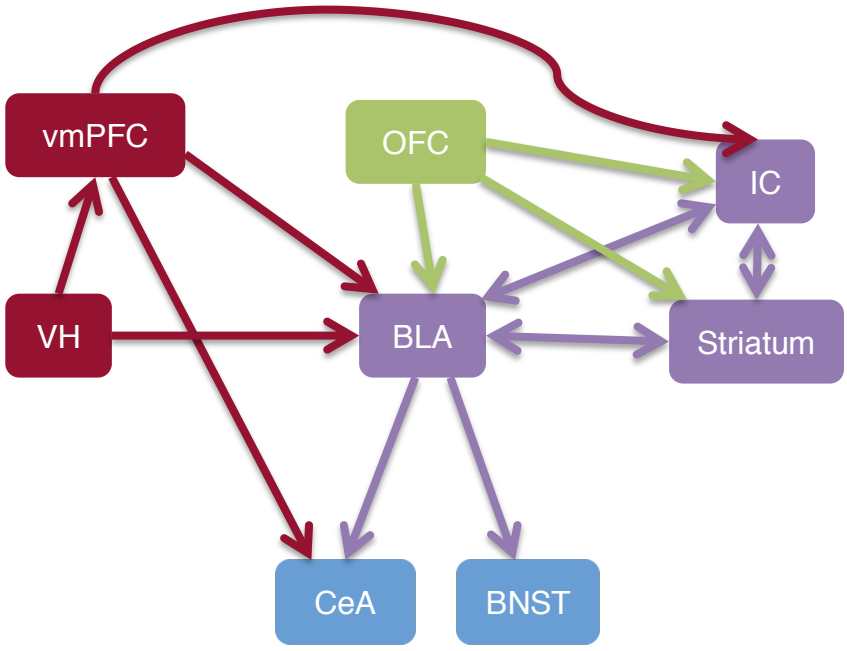

Figure 1.1 A hypothetical circuit for the processing of safety information. Red regions and arrows indicate a site of danger processing and the projection of danger information. Green regions and arrow indicate regions and transferring of safety information. Purple colored regions and arrows indicate regions that display altered patterns of responding due to the reception of safety information. Blue regions indicate regions that project to the regions, which ultimately lead to behavioral outputs.

Together, the work presented here begins to uncover the neural mechanisms that allow for modulation of fear in the presence of safety cues. By adding new nodes to a potential safety circuit, eliminating others, and narrowing in on the precise role of certain brain regions, this work adds to the growing literature on the neural basis of fear modulation. This research also adds to the growing field of sex as a biological factor, as sex differences in learning and memory and the mechanisms that underlie these behavioral differences are just beginning to be discovered. Importantly, the work here also creates new questions that will lead to precise, hypothesis-driven experiments for future work.

CHAPTER 2

Mechanistic Studies on the Acquisition and Recall of Fear Discrimination

The work in this chapter is published in the following manuscripts:

Sarlitto MC, Foilb AR, Christianson JP (2018) Inactivation of the Ventrolateral Orbitofrontal Cortex Impairs Flexible Use of Safety Signals. Neuroscience, 379:350–358.

Chen VM, Foilb AR, Christianson JP (2016) Inactivation of ventral hippocampus interfered with cued-fear acquisition but did not influence later recall or discrimination. Behav Brain Res, 296:249–253.

Foilb AR, Flyer-Adams JG, Maier SF, Christianson JP (2016) Posterior insular cortex is necessary for conditioned inhibition of fear. Neurobiol Learn Mem, 134 Pt B:317-27.*

**Portions of this work were submitted in partial fulfillment of the requirements for the Masters of Arts at Boston College in October 2016.*

2.1 Introduction

An animal's prosperity and survival require flexible adaptation to a constantly changing environment. Past experience shapes decision making in part through the accumulation of learned associations between stimuli and their associated outcomes, which can be used to make predictions and decisions about future behaviors. Yet the mechanisms by which explicit environmental cues come to predict safety are largely unknown (Christianson et al., 2012; Kong et al., 2014). Learned safety signals are potent modulators of behavior and have the ability to inhibit fear responses, such as behavioral freezing, and promote exploration or foraging when presented in compound with learned danger cues (Konorski, 1967; Myers and Davis, 2004; Rogan et al., 2005; Pollak et al., 2008; Christianson et al., 2011; Sangha et al., 2014; Chen et al., 2016). The neuroanatomical loci that mediate learning and recall of safety cues are largely unknown, but the vast wealth of knowledge about fear learning and discriminative learning can be used to inform investigations on the neurological correlates of safety/danger discrimination. In the research presented here, we explore the roles of orbitofrontal cortex, ventral hippocampus, and insular cortex in the acquisition and recall of safety signals.

2.1.1 Orbitofrontal Cortex

Areas in the frontal cortex are implicated in associative learning and, of these, the orbitofrontal cortex (OFC) seems to be important in mediating

cognitive flexibility (Dalley et al., 2004; Murray and Izquierdo, 2007; Rudebeck and Murray, 2014). There is a vast and rapidly growing body of literature in which the OFC appears to contribute to a wide array of cognitive functions, including novel stimulus-outcome contingencies, inhibitory control of appetitive learning, reward valuation, reversal learning and extinction (for reviews, see Stalnaker et al., 2015; Izquierdo, 2017). The 'cognitive map' hypothesis accounts for much of the empirical data regarding OFC function (Wilson et al., 2014). A cognitive map is the mental representation of the external environment (Gallistel, 1989). This theory proposes that the OFC maintains a cognitive map characterized by the current task state. For example, using a Pavlovian conditioning paradigm, a rat might be trained with repeated cue presentations to associate a conditioned stimulus (CS, like a flashing light) with the occurrence of an unconditioned stimulus reward (US, food pellet). As the rat begins to behaviorally distinguish between CS/no-CS trials, a reflection of differential reward expectancy, the cognitive map theory suggests that the two conditions become encoded as separate 'task states'. Tracking task states, particularly those occurring in the same context but with opposing outcomes, is considered by this theory to be a key function of OFC activity (Wilson et al., 2014). Thus, loss of OFC function does not always impair state-relevant processing but does invoke noticeable impairments to performance of tasks requiring differentiation between perceptually similar states.

The cognitive map model captures a significant amount of data relating to the function of OFC but a limitation of the theory stems from the nearly exclusive reliance on behavioral tasks in which the outcome expectancies related to the value of desirable stimuli, such as food. Regarding aversively motivated behaviors, such as conditioned fear in which several contributions of the OFC are reported (*e.g.,* Rodriguez-Romaguera et al., 2015; Zimmerman et al., 2018) but there are conflicting results (Shiba, Santangelo, & Roberts, 2016). A test of OFC function in a learning context in which the conditioned stimuli become associated with different expectations of aversive outcomes would provide a systematic test of the current theory's generality and provide insight to the neural circuitry underlying cognitive flexibility in the face of environmental stimuli that predict danger. Discrimination between danger and safety cues leads to differential freezing upon presentations of conditioned danger (shock paired A+) and conditioned safety (unpaired B-) cues. In the fear discrimination learning scenario, the A and B cues could be argued to reflect distinguishable task states that can be called upon to guide behavioral responses, and therefore involve the OFC.

The OFC receives polymodal sensory input and projects extensively to the limbic system and midbrain motor regions capable of governing behavioral output (Paxinos and Watson, 2007). It is this intersection of polymodal sensory input and visceromotor outputs that makes the OFC relevant to reward processing but also positions it well to map the outcome expectancies of undesirable aversive

stimuli. The rat OFC is subdivided into the medial (MO), ventral (VO), ventrolateral (VLO) and lateral (LO) and anterior agranular insular cortex (Price, 2007). Anatomical, mechanistic and recording experiments targeting these regions provide evidence for some regional functional specificity with regard to decision-making, with medial regions relating more to affective regulation with lateral regions more to sensory integration (Rempel-Clower, 2007; Izquierdo, 2017).

2.1.2 Hippocampus

Fear is also an important component of safety learning and fear responses are typically specific to contexts or stimuli that were previously paired with aversive stimulation (Maren et al., 2013). A growing number of studies implicate the ventral hippocampus (VH) in the acquisition of conditioned fear and the VH contributes to the modulation of fear in extinction and discrimination processes (Richmond et al., 1999; Bast et al., 2001; Zhang et al., 2001; Esclassan et al., 2009; Czerniawski et al., 2012; Wang et al., 2012; Cox et al., 2013; Zhang et al., 2014). The extinction of fear conditioned to a discrete cue depends upon the context in which extinction occurred because when the fear cue is presented in a novel setting, the fear response returns (Bouton and King, 1983). The return of fear in a novel context depends upon excitatory input from the VH to the amygdala and prefrontal cortex (Orsini et al., 2011). Thus, the VH contributes to both the initial acquisition of fear and the conditional expression of fear when

context is a discriminate feature (Hobin et al., 2006). Regarding discrete cues, hippocampal lesions interfere with the recall of a feature negative discrimination (Heldt et al., 2002), but the VH was not isolated in this study. Based on the role of VH in fear and the modulation of fear, it is possible that this region also plays a critical role in the reduction of fear to a safety cue.

2.1.3 Insular Cortex

When a safety cue is well-learned, it can become a conditioned fear inhibitor. In conditioned inhibition, fear is blunted by safety cues (Konorski, 1948; Rescorla, 1969; Christianson et al., 2012). A number of important investigations of conditioned inhibition in the context of appetitive learning identified the ventromedial prefrontal cortex (Rhodes and Killcross, 2007; MacLeod and Bucci, 2010; Meyer and Bucci, 2014), retrosplenial cortex (Robinson et al., 2011), central nucleus of the amygdala (Holland, 2012), the perirhinal and postrhinal cortex (Campolattaro and Freeman, 2006; Gastelum et al., 2012) and the serotonin system (Lister et al., 1996; Watkins et al., 1998). However, there are discrepancies between the appetitive and fear literatures with many more null results reported with regard to mechanisms for conditioned fear inhibition.

Insular cortex (IC) has a number of features that position it to contribute to identifying environmental safety cues. These include access to auditory, visual, and somatosensory information (Sudakov et al., 1971; Robinson and Burton, 1980a, b, c; Mufson and Mesulam, 1982; Shi and Cassell, 1998a, b; Remple et

al., 2003; Benison et al., 2007; Gogolla et al., 2014; Gogolla, 2017), a somatotopically organized body representation (Benison et al., 2007), multisensory integration (Rodgers et al., 2008), afferent intracortical and thalamocortical connectivity and efferent amygdala projections (Shi and Cassell, 1998a, b; McDonald et al., 1999). As described in Chapter 1, numerous reports identify electrophysiological signatures and molecular correlates of safety signals the amygdala (Campeau et al., 1997; Rogan et al., 2005; Pollak et al., 2008; Sangha et al., 2013; Likhtik et al., 2014). Because IC connectivity to the amygdala varies across its length (Shi and Cassell, 1998a, b; McDonald et al., 1999) and others have reported roles for anterior divisions of rodent insula in fear (Bermudez-Rattoni, 2014; Casanova et al., 2016), we targeted three points along the rostro-caudal axis. Extant data suggesting a role for posterior IC (pIC) in processing safety signals were obtained using a backwards conditioned safety signal in the context of an unpredictable traumatic stressor. In the midst of the traumatic stressor, safety signals prevented the development of numerous stressor sequellae; these “safety signal effects” were blocked by both lesion and pharmacological inactivation of the pIC (Christianson et al., 2008, 2011). We aimed to uncover whether the pIC contributes specifically to the learning or recall of safety signals in a fear discrimination paradigm.

2.1.4 Chapter Aims

In the studies presented here, we examined three brain regions—ventrolateral OFC (vlOFC), ventral hippocampus (VH), and insular cortex (IC)—in the learning and recall of safety cues in a conditioned fear discrimination paradigm. We tested if the behavioral expression of fear is under the flexible control of the danger or safe signals. This behavioral flexibility should entail the formation and use of a cognitive map and so provides a paradigm suitable to assess the generality of this role of the OFC in both learning and flexible control of behavior by aversive cues. We targeted the LO/VO regions with pharmacological inactivation in fear discrimination acquisition or later recall. Based on the role of VH in fear modulation, we hypothesized that the VH may play a general role in conditioned fear discrimination by using a fear discrimination paradigm where fear expression is controlled by the conditioned cue instead of a context. pIC was the focus of IC investigation due its previously discovered role in mitigation of stress by safety signals, and its known anatomical connectivity (Christianson et al., 2008, 2011; Shi and Cassell, 1998a, b). From this, we hypothesized that pIC was a likely site of integration for danger and safety information.

To examine the role of these brain regions in acquisition and recall of fear discrimination, as well as conditioned inhibition of fear, we used AX+/BX- fear discrimination conditioning, where animals were exposed to two CSs: a safe CS that was never paired with footshock (B) and a danger CS that was always paired

with shock (A). Brain regions were temporarily inactivated before either conditioning or recall testing and freezing was recorded as a behavioral measure of fear. We found that each of these regions plays a specific role in fear discrimination. vlOFC is specifically involved in recall of safe cues, VH appears to be involved in fear but not safety learning, and pIC, but not aIC or mIC, is specifically involved in the acquisition of a conditioned inhibitor, not in recall or acquisition of fear discrimination. The roles of vlOFC and pIC likely fit into a larger safety learning circuitry, which is the overarching aim of the work of this dissertation.

2.2 Materials And Methods

2.2.1 Animals

For experiments on the role of vlOFC, adult male Sprague-Dawley rats were obtained from Taconic (Hudson, NY), while experiments on VH and IC used adult male Sprague-Dawley rats obtained from Charles River Labs (Wilmington, MA). All rats weighed 250-300 g upon arrival, kept on a 12-hour light/dark cycle with lights. Rats single housed after surgery and a piece of autoclaved manzanita wood was provided for enrichment (Rosenzweig and Bennett, 1996) in accordance with recommendations by the Boston College Institutional Animal Care and Use Committee (IACUC). All animals were given 7-10 days to acclimate to vivarium before any procedures and a minimum of 7 days to recover

after any surgical procedures. All experimental protocols were reviewed and approved by the Boston College Institutional Animal Care and Use Committee.

2.2.2 Apparatus

All behavioral conditioning occurred in the same conditioning chamber. Conditioned stimuli were delivered via a white LED array and a speaker mounted at the top of the chamber. The chamber was illuminated with 2 infrared LEDs arrays and overhead cameras with infrared passing filters were used to recorded behavior. Freezing was detected using ANY-Maze software.

2.2.3 AX+/BX- Fear Discrimination Conditioning

Adapted from (Myers and Davis, 2004) and used previously (Chen et al., 2016; Foilb and Christianson, 2016), conditioning sessions involved 15 presentations each of shock-paired (A+) or unpaired (B-) cues, for a total of 45 minutes per session. Each trial was signaled by a common element (X), a 5 s, 1 kHz tone (75 dB) immediately followed by a 15 s discrete auditory (white noise pips, duration = 10 ms, rate = 3 Hz, 75 dB) or visual (flashing LED light, 264.0 Lux, 20 ms on/off) CS. The aversive unconditioned stimulus (US) was a 500 ms footshock (1.2 mA) that co-terminated with the A cue, such that each animal received 15 shocks per conditioning session. These parameters were adopted

based on the results of a pilot experiment in which the conditioned inhibition of freezing was assessed after conditioning with either serial or compound transfer stimuli (as in the AX+/BX- protocol of Myers & Davis, 2004); the serial conditioning protocol resulted in robust and reproducible conditioned inhibition of freezing, while the compound cues did not (Foilb and Christianson, 2016). As in Myers and Davis (2004), the variation used here retains common element X in conditioning, which serves as a transfer stimulus on B trials. Trials were presented in a quasi-random order, so that no cue occurred more than twice in succession. There was a fixed 70 s inter-trial-interval. Assignment of the light or pip as A or B cues was counterbalanced in each experiment, and equally represented in each treatment condition.

2.2.4 Discrimination Recall Tests

Recall tests were performed in the same apparatus as conditioning. In studies on vlOFC, recall testing consisted of a 2 min exposure to the context followed by 10 A and 10 B cues (30 s duration) presented in quasi-random order with a 30s inter-cue-interval. In the IC experiments focused on fear discrimination and VH experiments, discrimination was assessed by presenting the A and B cues (60 s duration) 6 times each in a quasi-random order. These tests also began with 2 min of baseline context exposure. For the majority of figures and analysis, the presentations of each cue were averaged across the test session.

2.2.5 Summation Tests

The efficacy of the B cue as a safety signal to inhibit behavioral freezing was assessed in summation tests. Tests began with 2 minutes of baseline context exposure, followed by a minute of the A cue, a minute of A and B cues presented in compound (AB) and a minute of the B cue. The cues were repeated for the following serial sequence: Baseline, A, AB, B, A, AB, B. For figures and analysis, the two presentations of each cue were averaged.

2.2.6 Cannula Placement and Microinjections

Surgical procedures were conducted in accordance with Boston College IACUC. Stainless steel guide cannula were implanted bilaterally to target vlOFC (+3.2mm anterior/posterior, AP from bregma, ±2.2mm medial/lateral, ML from the midline, -3.4mm dorsal/ventral, DV from dura), VH (AP -5.8mm, ML ±5.2mm, DV -7.0mm, all measured from bregma), anterior IC (AP +2.7mm, ML ±3.9mm, DV -5.2mm from bregma), medial IC (AP +0.5mm, ML ±4.9mm, DV -6.2mm from bregma), or posterior IC (AP -1.8mm, ML ±6.5mm, DV -6.2mm from bregma). While IC can be subdivided into granular, dysgranular and agranular regions along its dorsal-ventral axis, cannula were targeted for the central agranular region and the injection volume (0.5μL) was selected to permit diffusion throughout the three regions. Cannula tips found within any of the three subdivisions were included, thus conclusions from these were not intended to be specific to any of the IC subregions.

A minimum of 7 days of recovery were allotted before behavioral testing, during which time rats were periodically handled and stylets were checked to ensure the cannula remained unobstructed. Microinjections were made by gently restraining the rat in a cloth towel and replacing the stylet with a microinjector protruding 1 mm beyond the cannula tip. At the end of each experiment, rats were overdosed with tribromoelthanol, brains were removed and flash-frozen in 2-methylbutane on dry ice, and stored at -80ºC until they were sliced at 40 µm on a freezing cryostat (-20ºC). Slices were stained with cresyl violet, coverslipped, and allowed to dry overnight before cannula placement was determined by comparison with the Rat Brain Atlas in Stereotaxic Coordinates (Paxinos and Watson, 2007). Data from rats for which cannulas were not found or were located outside of the targeted areas were excluded in statistical analysis.

2.2.7 Pharmacological Inactivation of Brain Regions

The $GABA_A$ agonist muscimol was used to temporarily inactivate brain regions in most of the experiments. Muscimol was dissolved in sterile saline at 100 ng/µL (for vlOFC and IC, Moscarello and LeDoux, 2013) or 500 ng/µL (for VH, Hobin et al., 2006). In IC experiments, NMDArs were blocked with receptor antagonist D-(-)-2-Amino-5-phosphonopentanoic acid (AP5). AP5 was dissolved in sterile saline at 6 µg/µL (as in Bast et al., 2005; Amat et al., 2014; Christianson et al., 2014). In each case, the drug was administered bilaterally at 0.5 µL per side at a rate of 1 µL/min, with an additional minute allowed for diffusion. Vehicle

treated animals received saline injections at the same volume and rate as the drug infusions. AP5 injections were completed 15 minutes before conditioning (Bast et al., 2005) and muscimol injections were completed an hour before testing (Amat et al., 2005, 2014).

2.2.8 Data Analysis

Freezing was analyzed as percent time freezing during the relative cue condition. In discrimination tests, a discrimination index was calculated as freezing to B divided by freezing to A times 100, so that a value of 100 reflected no discrimination between A and B, and values less than 100 indicated reduced fear to B compared to A. In summation tests, a summation index was calculated as freezing to AB divided by freezing to A times 100. Thus values greater than 100 would reflect excitatory summation whereas values less than 100 would reflect conditioned inhibition. Group differences in behavioral freezing data were then evaluated by analyses of variance (ANOVA) with drug treatment treated as a between-subjects factor, and cue, day or test treated as within-subjects factors, except where noted. Main effects and interactions were deemed significant with $p < 0.05$ and between-subjects post hoc comparisons were made with Tukey's HSD correction or Sidak's correction. All analyses were made using GraphPad Prism. Detailed statistics will not be presented here, but are available in the published manuscripts listed at the beginning of this chapter.

2.3 Summary Of Experiments And Results

2.3.1 Ventrolateral Orbitofrontal Cortex

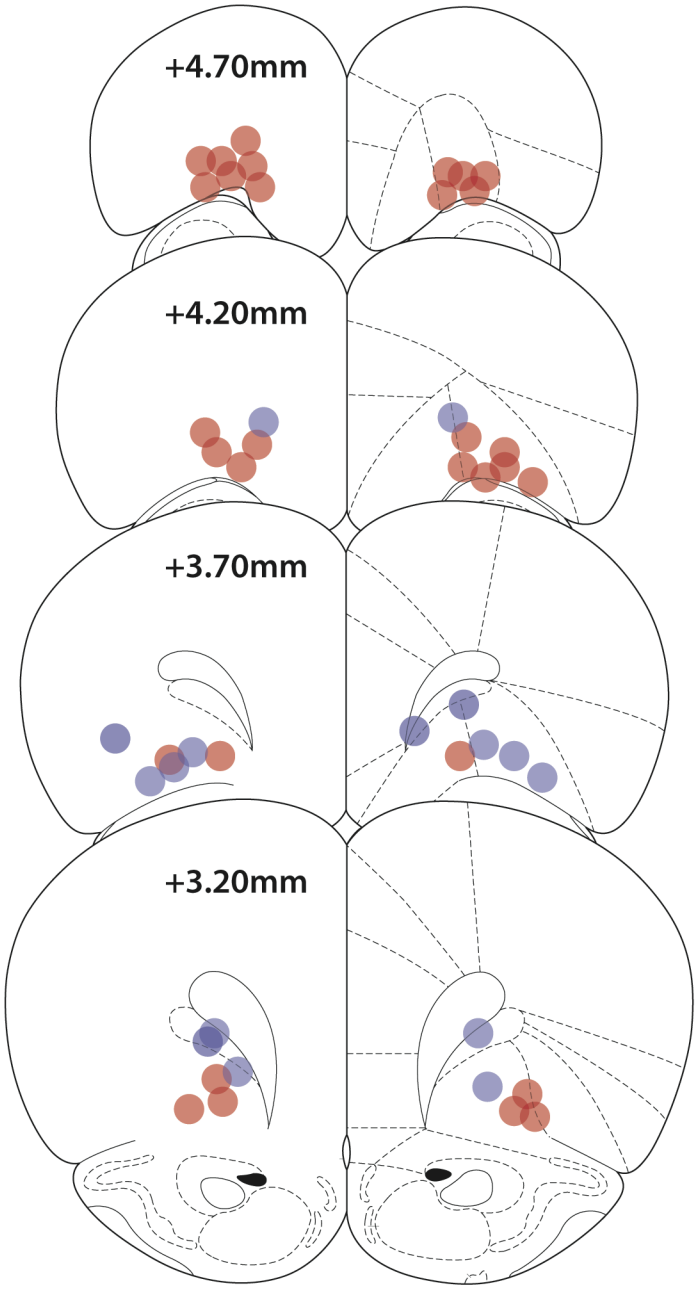

Figure 2.1 OFC cannula placements. Microinjection cannula tip locations included in Experiment 1.1 (blue circles) and Experiment 1.2 (red circles).

2.3.1.1.Experiment 1.1: Effect of pre-conditioning vlOFC inactivation on fear discrimination

To determine if the OFC contributes to the initial learning of the fear discrimination, 16 rats implanted with bilateral OFC cannula were given either muscimol or saline injections 1h prior to fear discrimination conditioning with the light and pip CSs. Behavioral freezing was quantified during the conditioning session and in a drug-free recall test 24 h later (cannula placements are depicted in Figure 2.1).

There was no significant main effect of drug on freezing during conditioning (Figure 2.2A, B), or during recall testing (Figure 2.2C, D). In the conditioning phase there was a significant main effect of trial and in the discrimination recall test there was a significant main effect of cue and a significant trial by cue interaction. Post hoc comparisons revealed significantly less freezing to B compared to A at each trial. To summarize discrimination behavior, and facilitate the display of individual subjects, a discrimination index was computed as a ratio of total freezing to B divided by freezing to A multiplied by 100. There was no difference in discrimination index between drug conditions during conditioning (Figure 2.2E) or recall (Figure 2.2F).

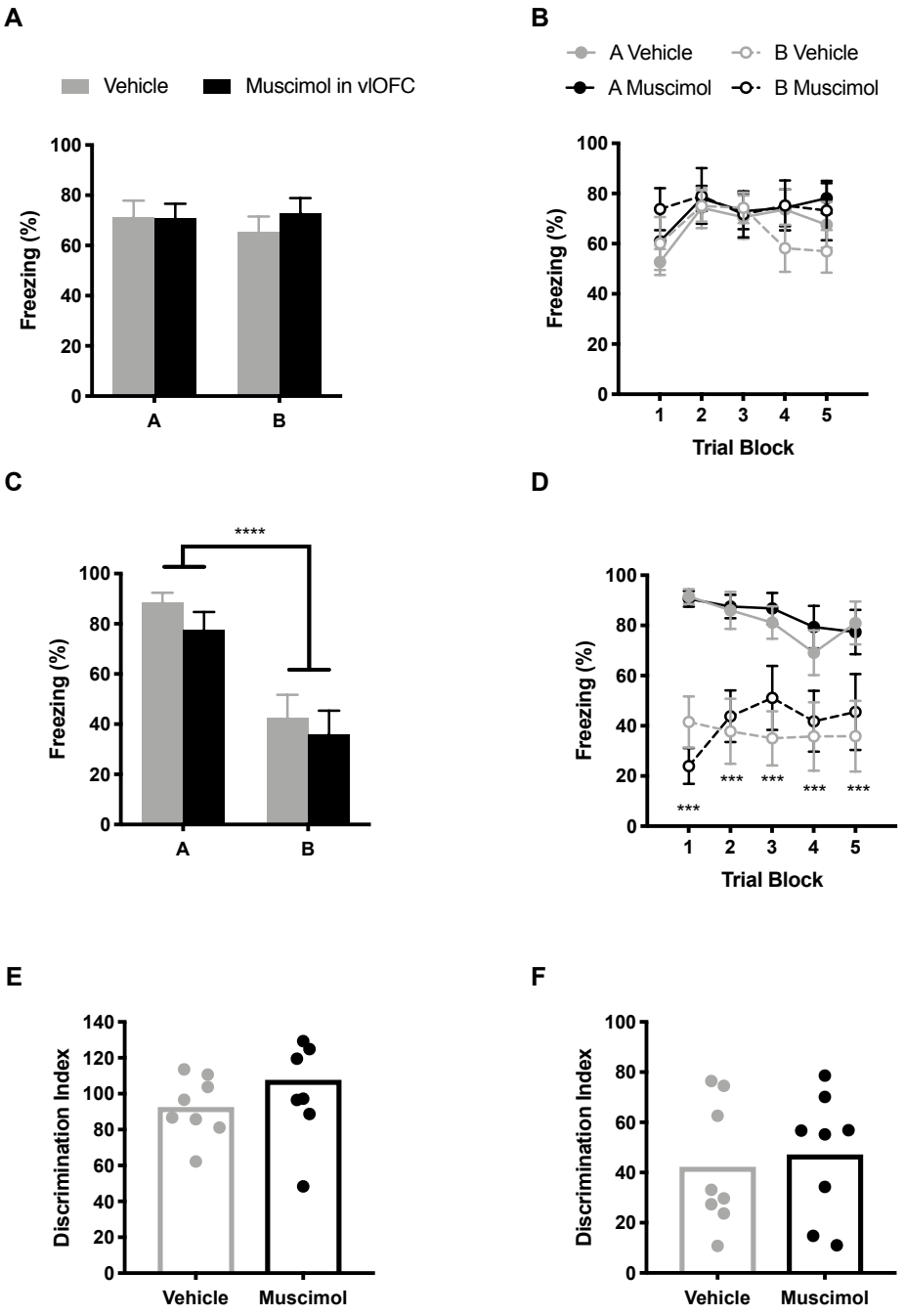

Figure 2.2 OFC in acquisition of fear discrimination. Rats were assigned to either intra-OFC muscimol or vehicle injections. **(A)** Mean (+SEM) freezing to the cues during conditioning and **(B)** Mean (±SEM) freezing to the cues in blocks of 3 conditioning trials. Pretreatment with muscimol did not alter freezing compared to vehicle controls. **(C)** Mean (+SEM) freezing to the cues in a recall test given 24h after discrimination conditioning and **(D)** Mean (±SEM) freezing to the cues in blocks of 2 trials during recall. There was no effect of pre-conditioning muscimol on later discrimination in the recall test. ***$p < 0.001$ freezing to B was significantly less than A in test average and across all trial blocks. **(E)** Mean (individual replicates) discrimination index (freezing to B divided by A times 100) during conditioning and **(F)** during recall testing. Muscimol and vehicle treated rats behaved equivocally in the discrimination task at both time points.

2.3.1.2 Experiment 1.2: Effect of post-conditioning vlOFC inactivation on fear discrimination during recall

To determine if the OFC is involved in the behavioral responses to danger and safety, 16 rats implanted with bilateral OFC cannula were given 3 consecutive days of fear discrimination conditioning, with a recall test each following morning. Three conditioning sessions were used to ensure that all rats exhibited differential freezing to the A and B during recall tests. There was differential freezing to A and B cues in the first and second discrimination recall tests (drug-free), with main effects of trial and cue and significantly different freezing between A and B on all trial blocks (Figure 2.3A).

Twenty-four hours after the third conditioning session, one half of the rats received bilateral muscimol infusions 1h before a recall test and the other half received a saline injection. Behavioral freezing was quantified during the recall test and the rats were returned to their vivarium. To allow a within-subjects comparison, 24h later all rats received a second round of microinjections but drug treatment was switched such that rats that received muscimol in the first test, received saline in the second test and vice versa. Muscimol prior to recall testing appeared to interfere with fear discrimination especially in the later trials of each test (Figure 2.3B). Data from Test 3 and Test 4 were analyzed separately. In Test 3 there were significant main effects of drug, trial, and a significant cue by drug interaction. In Test 4 there were significant main effects of drug, cue, and a significant cue by trial interaction. In Tests 3 and 4, there was significantly less

freezing to B in the vehicle group compared to muscimol on Trials 4 and 5. Discrimination was evident in both drug groups with freezing to A differing from freezing to B on all trials in the vehicle condition in both Tests 3 and 4 and in the muscimol condition in Test 3 trials 1, 2, 3, and 4 and Test 4 trials 1, 2, 3, and 5.

The pattern of behavior in Tests 3 and 4 was consistent except for a slight reduction in overall freezing in Test 4 as seen in Figure 2.3B, therefore we pooled data from Tests 3 and 4, averaged freezing across trials, and freezing was analyzed (Figure 2.3C). There was a significant main effect of cue, drug and cue by drug interaction. The apparent discrimination impairment after muscimol could reflect a generalized effect of muscimol on freezing *per se.* To isolate the discrimination component, discrimination indices were computed for the pooled data in Figure 2.3D. The discrimination index was significantly greater in muscimol treated animals compared to vehicles, indicating weaker inhibition of freezing to B. In visual inspection of freezing during Tests 3 and 4 (Figure 2.3B) it appeared that in the muscimol treated group, freezing to B was initially similar to the vehicle group, but drifted toward A cue level fear over repeated trials. To quantify this trend, we computed discrimination indices for the first and last trial blocks (Figure 2.3E). The discrimination index was greatest in the muscimol trial 5 conditioning, which was greater than vehicle at both trial 1 and trial 5, but not from muscimol trial 1. However, the interaction of trial and drug did not reach significance in this comparison. The trend that OFC muscimol effects occurred primarily in the later test trials requires additional investigation.

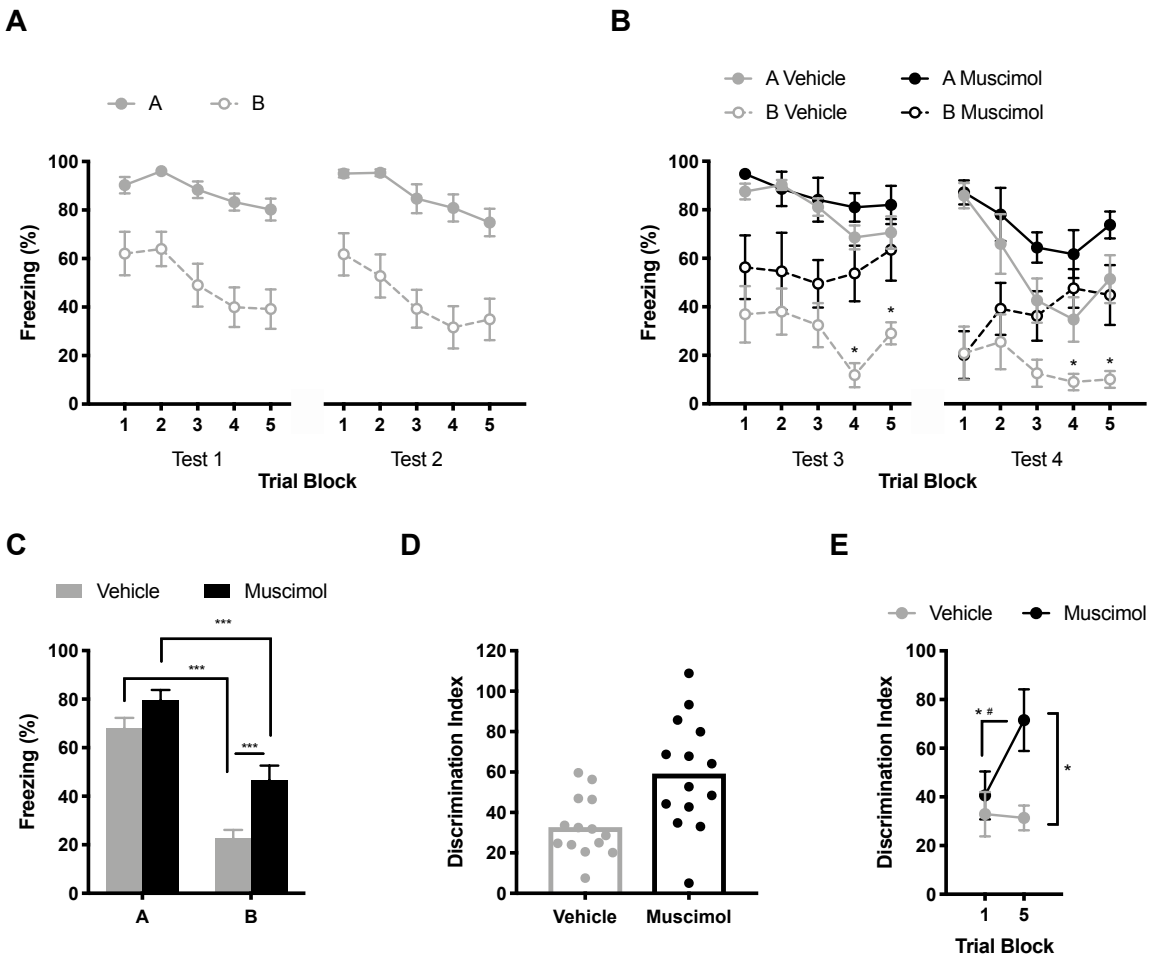

Figure 2.3 OFC in recall of fear discrimination. (A) Mean (±SEM) freezing to A and B in blocks of 2 discrimination recall trials in tests 1 and 2 prior to drug administration. Robust discrimination was evident with significantly reduced freezing to B at every trial, *p*s < 0.004. **(B)** Mean (±SEM) freezing to A and B in blocks of 2 trials 60 min after injection of muscimol or vehicle. Discrimination was evident in both tests, but greater discrimination was evident in the vehicle groups in the later trials. **p*s < 0.019 B vehicle vs. B muscimol. **(C)** Mean (+SEM) freezing to the A and B pooled across test days 3 and 4, and trials. Differential freezing was significant in both vehicle and muscimol conditions but there was greater freezing expressed to B in muscimol treated animals. ****p* < 0.001. **(D)** Mean (individual replicates) fear discrimination index where suppression of freezing to B was significantly better in the vehicle group ****p* < 0.001. **(E)** Mean fear discrimination index (±SEM) pooled across tests 3 and 4 for the first and last trial blocks. Discrimination is equal between vehicle and muscimol groups at the outset of the recall test in trial block 1, but discrimination becomes much worse in the muscimol condition by the end whereas it remains stable in the vehicle control group. **p* < 0.05 Muscimol Trial 5 vs. Vehicle Trial 1 and Vehicle Trial 5, $^{\#}p_{\text{uncorrected}}$ = 0.02 Muscimol trial 5 vs. Muscimol Trial 1.

2.3.2 Ventral Hippocampus

2.3.2.1 Experiment 2.1: Effect of pre-conditioning VH inactivation on fear discrimination

To explore the role of VH in acquisition of fear discrimination, rats were implanted with bilateral cannula (VH cannula placements are depicted in Figure 2.4) and randomly assigned to muscimol or vehicle conditions on day 1. Rats were injected, returned to the homecage, and 1 h later received AX+/BX- conditioning. Fear discrimination recall was assessed on Days 2 and 3 in identical tests.

During conditioning, rats spent the majority of time freezing, but there were no significant main effect of cue, drug, or cue by drug interaction (Figure 2.5A). Discrimination indices did not differ between vehicle and muscimol conditions (Figure 2.5B) In the recall test on day 2, rats with prior muscimol exhibited reduced freezing to all cues (A, B and context) compared to vehicle condition (Figure 2.5C). There were main effects of drug, cue and a significant drug by cue interaction. Freezing in the vehicle condition was significantly higher to all cues compared to the muscimol condition. In the vehicle condition, discrimination was evident as significantly increased freezing to A compared to either B or context alone, and freezing to B was significantly less than to context. The recall test was repeated on day 3 (Figure 2.5D) with main effects of drug, cue, and a drug by cue interaction. As on day 2, there was significantly greater freezing to all cues in the vehicle condition relative to pre-training muscimol and discrimination was evident

in the vehicle condition as significantly greater freezing during A than during either B or the context.

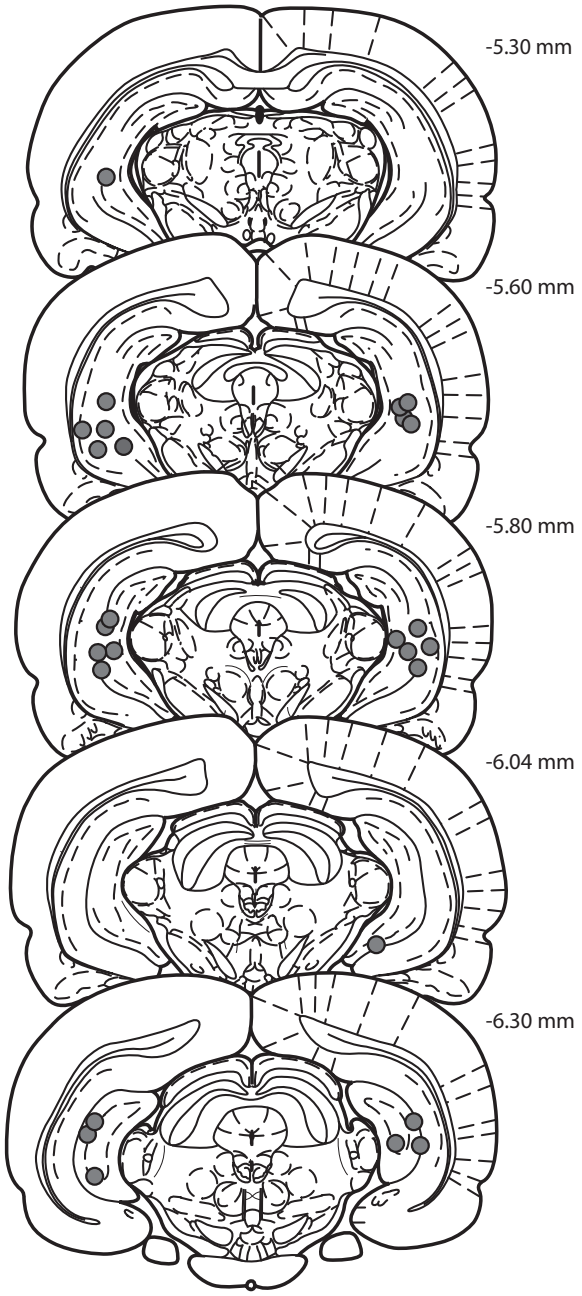

Figure 2.4 VH cannula placements. Microinjection cannula tip locations of rats included in Experiment 2.1 and Experiment 2.2.

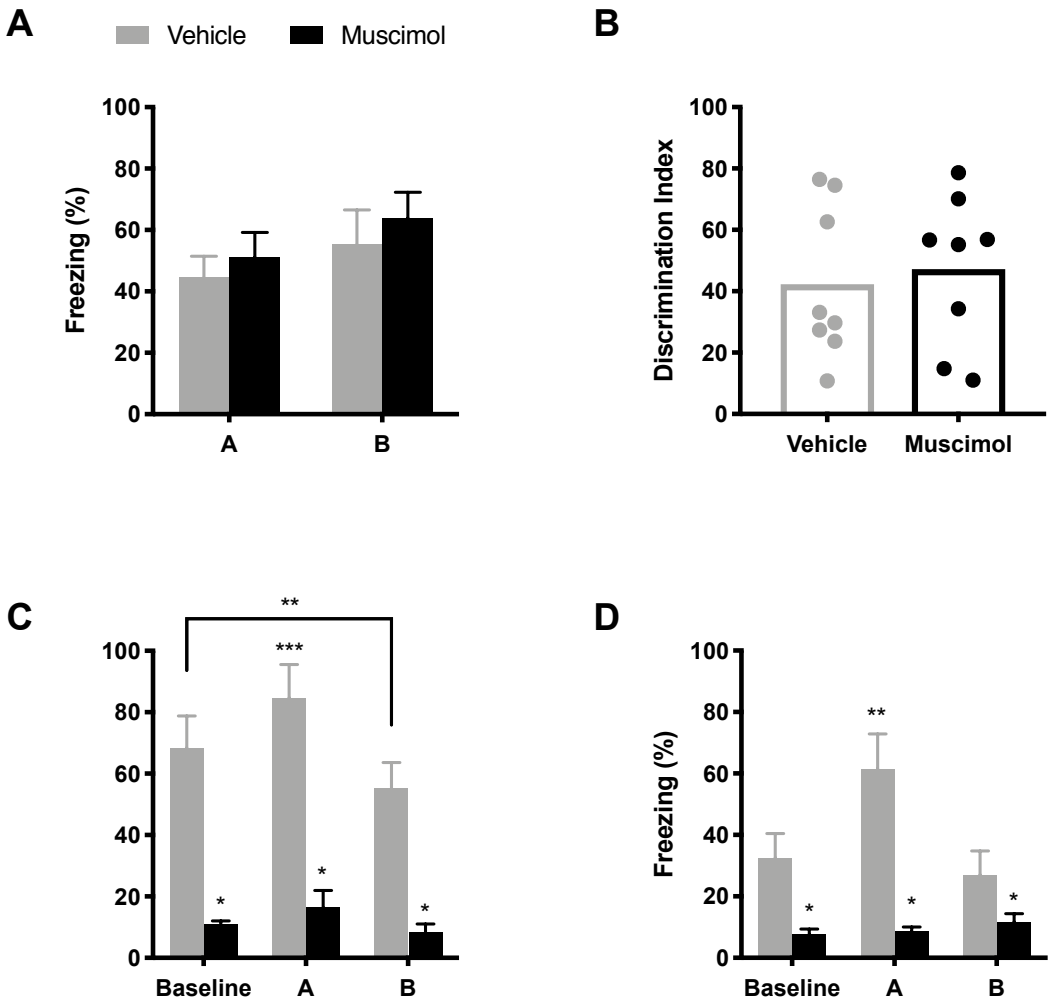

Figure 2.5 VH in acquisition of fear discrimination. (A) Mean (+SEM) freezing during A and B trials during conditioning. There was consistent freezing in all conditions with no effect of muscimol. **(B)** Mean (individual replicates) discrimination index indicated no differences between muscimol and vehicle treated animals during conditioning. **(C)** Mean (+SEM) freezing during recall on day 2. Rats with intra-VH vehicle injections prior to conditioning were able to discriminate between A and B and baseline context ($ps < 0.05$) and there was significantly less freezing to B than the baseline context ($p < 0.05$). In contrast, VH muscimol injections reduced freezing to all cues ($*ps < 0.01$). **(D)** Mean (+SEM) freezing during recall on day 3. Rats in the vehicle condition were able to discriminate between A and B cues ($ps < 0.05$) and rats with prior muscimol displayed reduced freezing to all cues ($*ps < 0.01$).

2.3.2.2 Experiment 2.2: Effect of post-conditioning VH inactivation on fear discrimination during recall

To test the role of VH in fear discrimination recall, all rats from Experiment 2.1 received additional conditioning and testing until both the vehicle and muscimol treated rats exhibited equal fear and discrimination. This required two additional drug-free AX+/BX- conditioning sessions, which began in the afternoon on Day 3 and again on Day 4. Recall tests were given on the morning of Day 4 and Day 5 at which point all rats exhibited equal freezing and discrimination, regardless of past drug treatment (Figure 2.6A). In this recall test, there was a significant main effect of cue, but no significant main effect of drug, or drug by cue interaction. Significantly greater freezing was displayed to A compared to B or context in each group, despite prior muscimol or vehicle treatment. Thus, prior to the final recall tests, the vehicle and muscimol treated rats exhibited equal fear recall and discrimination.

Rats were then assigned to new muscimol and vehicle groups each consisting of 4 rats from the previous muscimol group and 4 rats from the previous vehicle group. On day 8 rats received either muscimol or vehicle, according to their new groups and 1 h later given a final recall test. There was a significant main effect of cue, but no effects of drug or drug by cue interaction (Figure 2.6B). Discrimination was evident with significantly greater freezing to A compared to B, and to A compared to baseline context in both the muscimol and

vehicle conditions. Similarly, comparisons of the discrimination index during recall testing after muscimol or vehicle injections intra-VH did not significantly differ. This result indicates that VH is not necessary for the recall of fear discrimination.

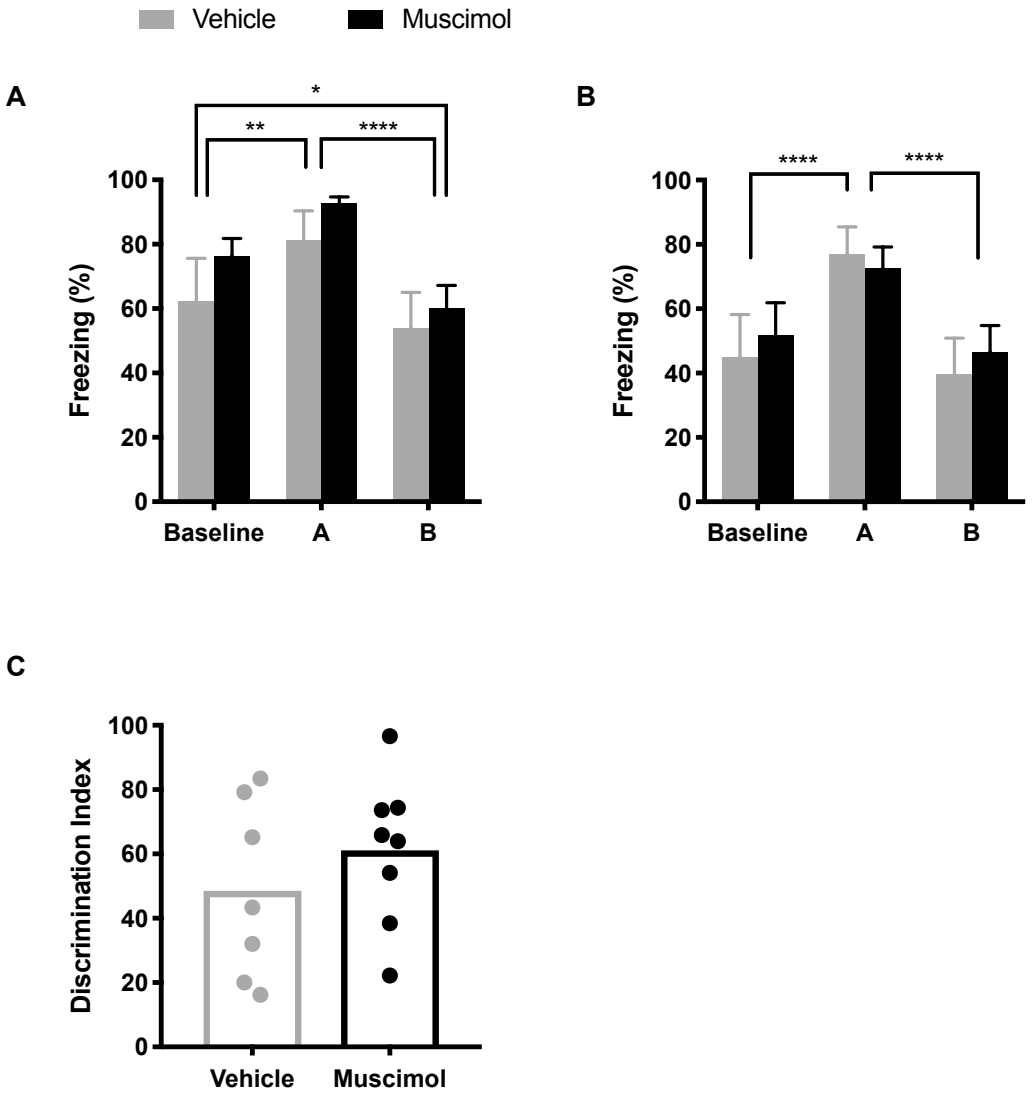

Figure 2.6 VH in recall of fear discrimination. (A) Mean (+SEM) freezing during recall on day 5 after reassignment to new muscimol and vehicle conditions. In both conditions, rats discriminated between A and B and baseline context (*ps* < 0.05). **(B)** Mean (+SEM) freezing during recall on day 8. Rats with vehicle injections or muscimol injections into the VH prior to the recall test were able to discriminate between A and B and baseline context (*p*s < 0.05). **(C)** Discrimination index (individual replicates) did not differ between muscimol- and vehicle-treated prior to recall testing.

2.3.3 Insular Cortex

2.3.3.1 Experiment 3.1: Insular cortex in acquisition of fear discrimination

To examine the involvement of IC in acquisition fear discrimination, rats received cannula implants within the pIC. AP5 was used to block N-methyl-D-aspartate receptors (NMDAr), which are critical to synaptic plasticity and numerous mnemonic functions (Morris, 2013). Cannula placements of all intra-IC injections are displayed in Figure 2.7. Injections of AP5 or saline to pIC were made 15 min before AX+/BX- conditioning. The following morning, animals were tested for fear discrimination.

There was no effect of AP5 on freezing during conditioning, with no main effect of drug or drug by cue interaction, but a significant main effect of cue (Figure 2.8A). Post hoc tests revealed the main effect of cue as differential freezing between baseline context and B, but no difference between A and B. In the discrimination test on the following day, rats in both AP5 and saline groups appeared to have acquired equal fear to A with discrimination evident as reduced freezing to B (Figure 2.8B). Accordingly, there was a main effect of cue, but no main effect of drug or interaction of drug and cue. Post hoc comparisons showed effective fear discrimination with significantly greater freezing to A compared to both B and baseline context. Therefore, pIC NMDAr do not appear to be critical to the acquisition of a basic fear discrimination

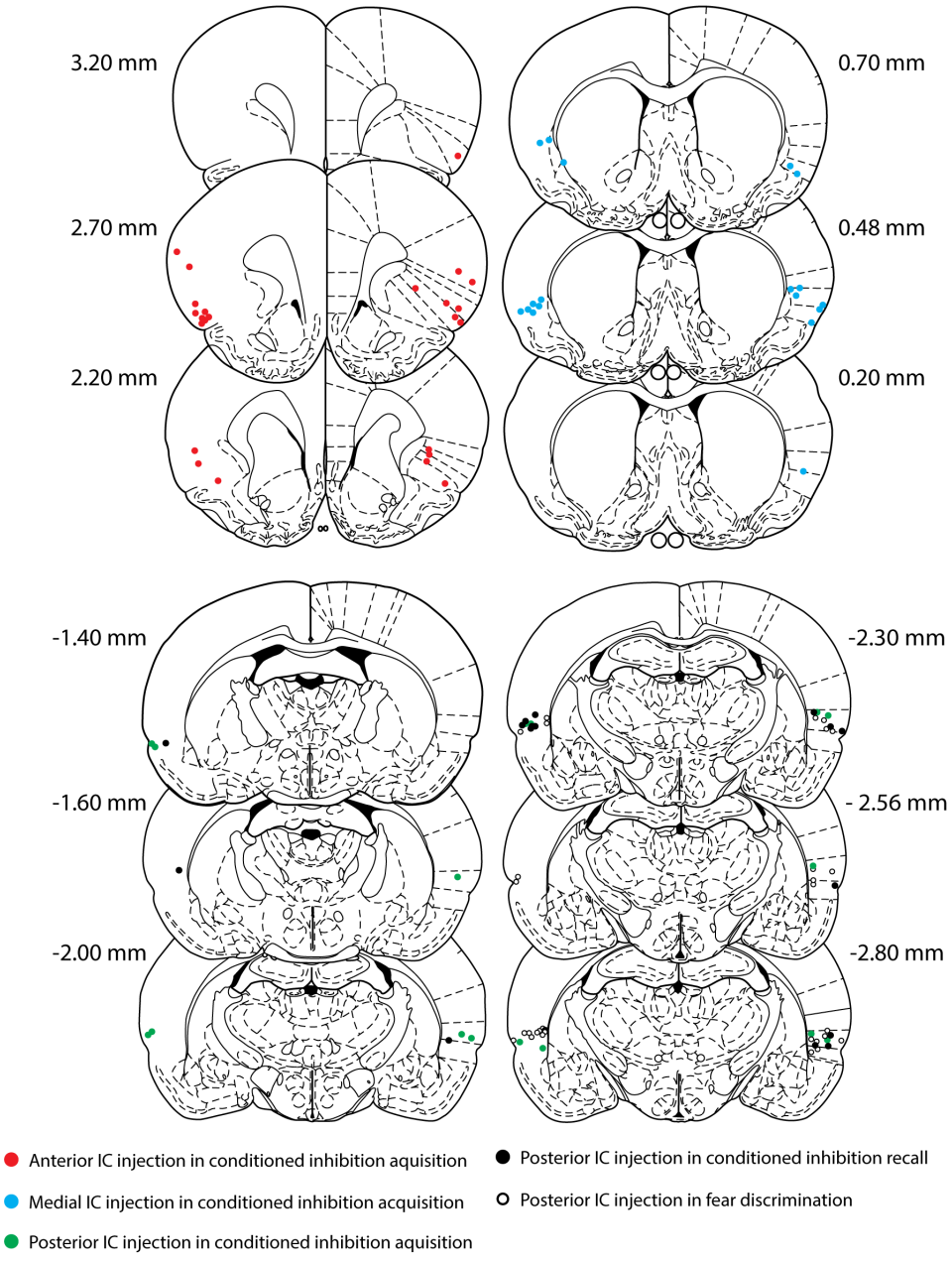

Figure 2.7 IC cannula placements. Sites of microinjections for all cannula experiments. White dots indicate posterior IC placements from Experiments 3.1 and 3.2. For Experiment 3.3, purple dots indicate anterior IC injections, grey dots indicate medial IC injections, and black dots are posterior IC injection sites. Blue dots indicate posterior IC injections in Experiment 3.4. Images were reconstructed from the atlas of Paxinos & Watson (1998).

.

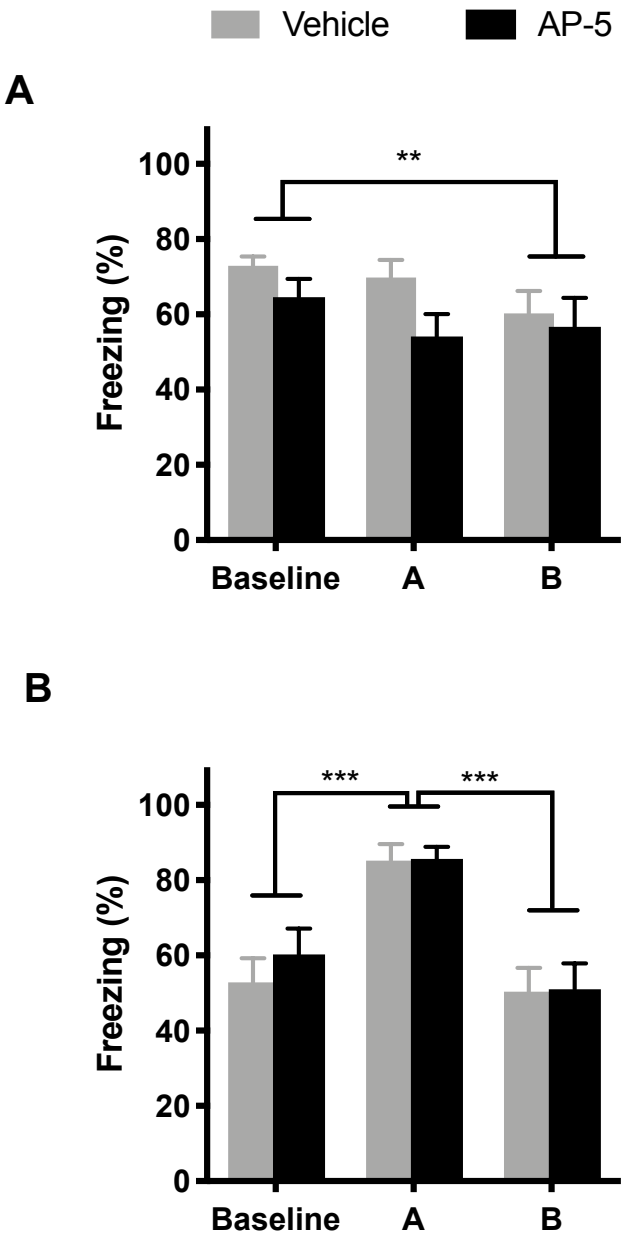

Figure 2.8 Posterior IC in acquisition of fear discrimination. (A) Mean (+SEM) percent freezing during conditioning after intra-IC AP5 injections. There was no effect of AP5, but significantly different freezing to baseline context and B cue presentation. **(B)** Mean (+SEM) percent freezing during recall test 1 after intra-IC AP5 injections before conditioning. There was no effect of drug, but a significant effect of cue, with increased freezing to A compared to both B and baseline context in both treatment conditions. Overhead brackets and asterisks indicate significant differences: ** $p < 0.01$, *** $p < 0.001$.

2.3.3.2 Experiment 2: Insular cortex in recall of fear discrimination

We next sought to determine if pIC neuronal activity was required for fear discrimination recall. Using the same set of animals, all rats received a second, drug-free conditioning session. Prior to a second fear discrimination recall test, muscimol or saline microinjections were made to pIC and fear discrimination was tested 1 hour later. To increase the experimental power, rats were given a third fear discrimination test and received the opposite muscimol or vehicle treatment 2 days later, for a within subjects comparison. On day 3, one half of the subjects received muscimol prior to the test while the other received vehicle; on day 5 the treatments were reversed. Rats were left alone for 1 day between tests to ensure washout of muscimol.

No effects of prior AP5 treatment were apparent in analyses so the data were pooled. Pretest muscimol did not appear to influence freezing to any cue or discrimination (Figure 2.9A). A significant main effect of cue, but no effect of drug, and no interaction of drug and cue were found. Freezing to each cue was significantly different from all other cues: baseline context vs. A, A vs. B, and baseline vs. B. Similarly, the discrimination index did not differ between muscimol and vehicle conditions (Figure 2.9B). These results indicate that pIC is not necessary for acquisition or recall of fear discrimination.

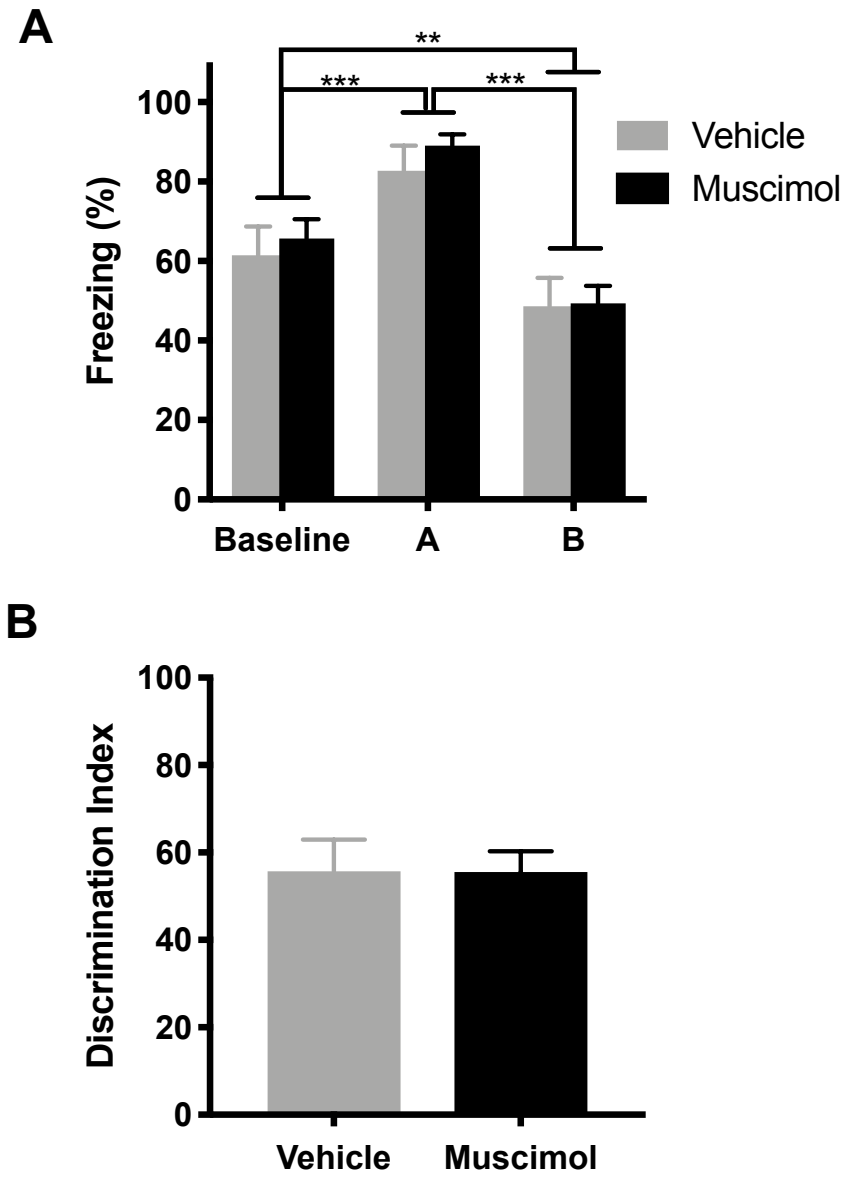

Figure 2.9 Posterior IC in recall of fear discrimination (A) Mean (+SEM) percent freezing in discrimination recall test 1 h after intra-IC muscimol injections. There was no effect of drug, but a main effect of cue where freezing to each cue was significantly different from all other cues. **(B)** Discrimination index, calculated as freezing to B divided by freezing to A multipled by 100, was not different in muscimol vs. vehicle treated animals. Overhead brackets and asterisks indicate significant differences as follows: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

2.3.3.3 Experiment 3.3: Role of insular cortex in the acquisition of conditioned inhibition of fear.

To test the necessity of IC for the acquisition of conditioned inhibition, cannulas were implanted in three regions of IC—anterior (aIC), medial (mIC), and posterior (pIC; Figure 2.7). Rats received conditioning for 5 days, with summation tests in the conditioning chamber each subsequent morning. In pilot experiments, rats that received injections of vehicle on the first day of conditioning never expressed conditioned inhibition, possibly due to tissue damage caused by repeated microinjections on 5 consecutive days. Since conditioned inhibition is not evident before day 3 (as shown in Foilb and Christianson, 2016), drug manipulations began prior to conditioning session 3. On days 3, 4, and 5, rats received intra-IC (aIC, mIC, or pIC) injections of AP5 or vehicle 15 minutes before conditioning (timeline displayed in Figure 2.10A). Only pIC NMDAr blockade interfered with acquisition of conditioned inhibition. During conditioning on day 5, animals with IC AP5 injections froze significantly more than vehicle or medial IC AP5 animals. This is consistent with the failure to inhibit fear in the presence of B in summation tests in this group and does not reflect a general effect of AP5 on fear expression. Performance in the summation tests for each drug and region group are shown in Figure 2.10B.

There were significant main effects of test and region. Post hoc analyses showed summation in vehicle animals was significantly reduced in tests 3, 4 and 5 compared to test 1, in tests 4 and 5 compared to test 2, and in test 5 compared

to test 3. This gradual decrease in summation index (indicating improved conditioned inhibition) was comparable to the animals without cannula or injections. Similarly, in animals with medial and anterior IC injections of AP5, summation scores were significantly lower in tests 4 and 5 compared to test 1, and further improved in test 5, which was significantly reduced compared to tests 2 and 3. Conversely, AP5 injections to the pIC resulted in significantly greater summation scores compared to aIC and mIC AP5 injections and vehicle injections on tests 4 and 5. Thus, only pIC NMDAr appear to be critical to the acquisition of conditioned inhibition as measured by summation.

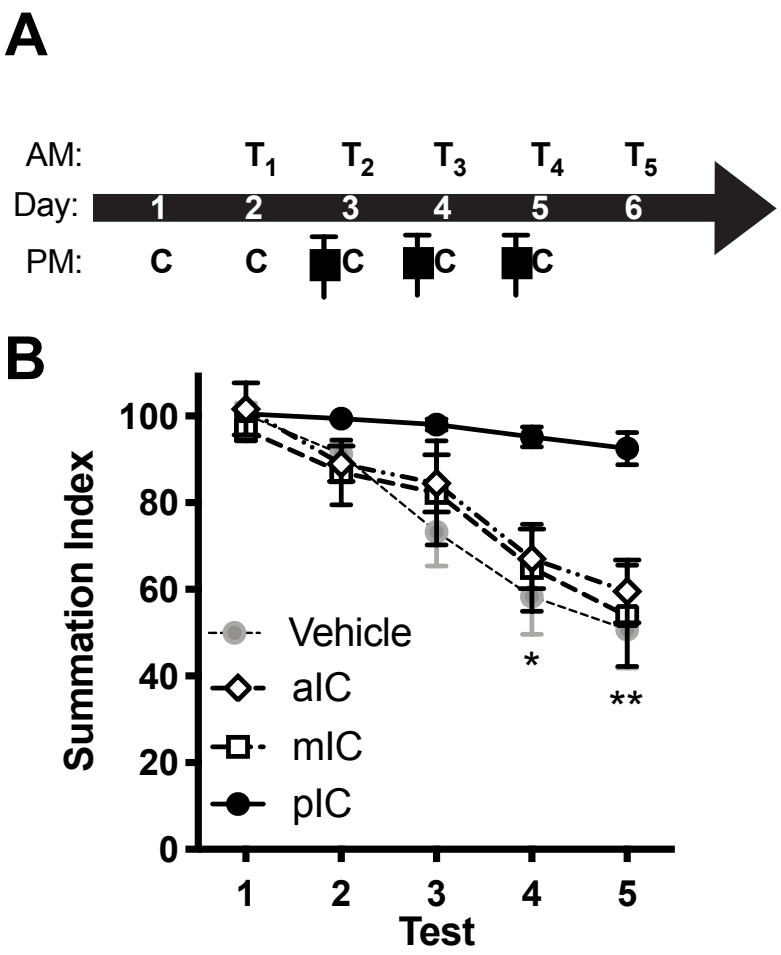

Figure 2.10 IC in the acquisition of conditioned inhibition of fear. (A) Timeline of conditioning, C, summation tests, T, and intra-IC AP5 injections (syringe icon) before conditioning. **(B)** Mean (±SEM) summation index in tests given the morning after pre-conditioning with vehicle or AP5 injections to anterior (aIC), medial (mIC) or posterior IC (pIC) on days 3, 4, and 5. Summation index was calculated as freezing to AB divided by freezing to A times 100. Greater conditioned inhibition was evident as lower summation scores in all groups compared to the pIC in tests 4 and 5. All groups except the pIC showed significant improvement in summation over the course of conditioning with significantly lower summation indices in tests 4 and 5 compared to their respective tests 1, 2 and 3 (ps < 0.05).

2.3.3.4 Experiment 3.4: Role of insular cortex in the recall of conditioned inhibition of fear.

Because the B cue appears to gain strength as an inhibitor after each conditioning session, it is possible that recall and reconsolidation processes occur within the pIC during the conditioning sessions. Thus, the preceding results could be attributed to a role of pIC in either the recall of B or the subsequent consolidation of new learning to B during conditioning. To investigate this possibility that pIC contributes to the recall of either A or B cues and to test whether pIC contributes to the expression of summation, a separate set of rats were implanted with cannula in the pIC (Figure 2.7) and received conditioning on consecutive days until conditioned inhibition was evident in summation tests given each subsequent morning. Importantly, animals were handled 1 h prior to each summation test to habituate to the microinjection procedure (timeline displayed in Figure 2.11A). Handling prior to summation tests delayed the acquisition of conditioned inhibition; therefore rats were only included in the inactivation phase of the experiment if they reached a summation index of 80 or less (as in Likhtik et al., 2014). Before the final summation test, pIC neuronal activity was pharmacologically inhibited by intra-IC injections of the $GABA_A$ receptor agonist muscimol or vehicle saline 1 h prior to a final summation test.

Pharmacological inactivation of pIC before recall surprisingly reduced freezing to all cues (Figure 2.11B). There were significant main effects of drug and cue, and because there was no significant drug by cue interaction, post hoc

analyses included both groups. Although freezing was reduced, all rats appeared to discriminate between the different cues in the summation test. Conditioned inhibition remained intact with significantly reduced freezing to baseline context, AB and B compared to A. The main effect of drug was evident as a significant reduction in fear in the muscimol condition compared to vehicles, where muscimol-treated animals froze significantly less to presentations of A and AB. The results of IC studies, together, implicate pIC specifically in the acquisition of conditioned inhibition of fear, rather than recall or prerequisite fear discrimination.

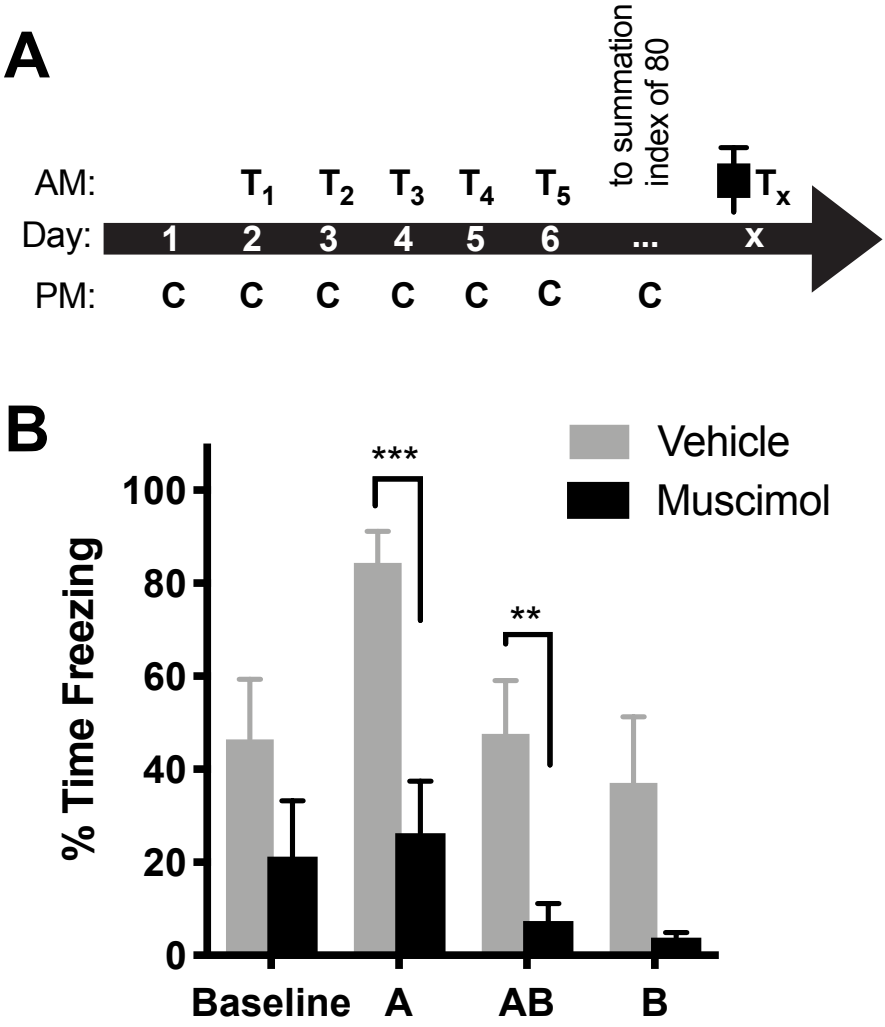

Figure 2.11 pIC in recall of conditioned inhibition of fear. (A) Timeline of conditioning, summation tests, and muscimol injections to the posterior IC before a summation test. Animals received repeated conditioning and recall testing until they reached a summation index less than 80. **(B)** Mean (+SEM) freezing to baseline context, A, AB and B 1h after muscimol injections. There were main effects of both drug and cue where animals froze significantly more to A than to the AB or B and there was significantly reduced freezing to presentations of A and AB in the muscimol-treated animals compared to vehicles. Overhead brackets and asterisks indicate significant differences, ** $p < 0.01$, *** $p < 0.001$.

2.4 Discussion

The goal of these projects was to advance understanding of the neural circuitry that mediates discrimination between safe and dangerous cues and fear inhibition by learned safety signals. To do this, we used AX+/BX- fear discrimination conditioning and region specific inhibition or blockade of NMDAr to test the necessities of vlOFC, VH, and IC in acquisition and recall of fear discrimination. This collection of studies provides mechanistic evidence for the involvement of vlOFC in recall of safe cues, necessity of pIC in acquisition of conditioned inhibition of fear, and role of VH in the acquisition of fear, but not in discrimination or safety learning. Here I review some of the implications of these findings; more detailed discussions can be found in the original publications.

2.4.1 Orbitofrontal Cortex

Pharmacological inactivation of the VO/LO region of the OFC prior to conditioning did not appear to influence any aspect of the fear discrimination conditioning, or the later flexible response to the cues. In contrast, when vlOFC was inactivated prior to a recall test, discrimination was impaired. The impairment was more prominent in the later test trials in which freezing responses to the B cue increased, suggesting that OFC is not involved in the initial recall of the discrimination. These results are generally consistent with the literature implicating the OFC and its outputs in tasks requiring cognitive flexibility.

Importantly, the demonstration of a role for OFC in flexible responses in an aversively motivated learning paradigm suggests that the “cognitive map” theory of OFC function generalizes to many types of learning. Although the OFC contributes to attention and salience assignment (Kahnt and Tobler, 2013; Ogawa et al., 2013) which would suggest a role for the OFC in the initial danger learning and discrimination, there are a number of reports where OFC lesions or inactivation did not consistently interfere with simple Pavlovian conditioning (Gallagher et al., 1999; McDannald et al., 2005; Ostlund and Balleine, 2007; Gremel and Costa, 2013).

In Experiment 1.2, pharmacological inhibition of the OFC both increased fear expression and interfered with behavioral flexibility in recall testing (Figure 2.3). Recent work also found that OFC lesions resulted in increased fear generalization in a danger uncertainty task (Ray et al., 2018), providing further evidence that the may OFC play a general role in the inhibition of freezing or that it tracks the changing expectations of danger. Interference with either function would manifest as a deficit in fear discrimination. Interestingly, OFC inhibition did not influence overall fear levels or fear discrimination in the first few safety trials. OFC inhibition appeared to bias the behavioral response towards danger later in the test, at which point the animals had toggled between relatively high and low freezing levels several times. This is consistent with procedurally different, but conceptually related, reports in which OFC inhibition with muscimol (Clark et al., 2008) or chemogenetic inhibition (Zimmermann et al., 2018) did not acutely alter

aversively motivated behavior (punishment and conditioned fear, respectively), but did increase the influence of the aversive CS on behavior in later tests (i.e. fear extinction recall). In appetitive studies, OFC has also been found specifically necessary for stable, well-learned stimulus-outcome contingencies (Riceberg and Shapiro, 2012; Izquierdo, 2017; Riceberg and Shapiro, 2017). The results of the experiments on vlOFC in fear discrimination reinforce a prevailing view of OFC function in the flexible use of learned associations to modulate behavioral responses.

2.4.2 Hippocampus

Our experiments on VH added to the growing number of studies that investigated the VH in fear learning and recall. Pre-training infusion of muscimol into the VH impaired the acquisition of fear learning to all stimuli present during conditioning. In contrast, vehicle treated rats were able to discriminate between the cues with reduced freezing to B compared to A. The same pattern occurred in a second recall test at 48h post-training which suggests that lingering effects of muscimol did not influence recall. Pre-testing inhibition of the VH did not alter freezing to A and B compared to vehicle animals.

Considering fear conditioned to discrete cues, our results in VH corroborate a large number of studies implicating VH in fear acquisition (Richmond et al., 1999; Bast et al., 2001; Zhang et al., 2001; Esclassan et al., 2009; Czerniawski et al., 2012; Wang et al., 2012; Cox et al., 2013; Zhang et al.,

2014). Concerning recall, however, VH inhibition sometimes has no effect and other times causes a significant reduction in fear (Maren and Holt, 2004; Sierra-Mercado et al., 2011). While our results do not preclude the possibility that a circuit including the VH would contribute in some way to conditioned discrimination acquisition, it does call into question the necessity of the VH in learned discrimination. This interpretation is consonant with a recent report that CS evoked hippocampus potentials did not distinguish between auditory danger and safety stimuli (Likhtik et al., 2014). The current results indicate a more general role of VH in a fear circuit but VH itself may not encode excitatory and inhibitory associations to discrete CSs. Additional research is warranted to determine what role, if any, the VH contributes to distinguishing between reinforced and unreinforced fear stimuli during conditioning.

2.4.3 Insular Cortex

Prior work had identified the pIC as a region capable of integrating the sensory information required to distinguish between safety and danger and modulate the output of fear circuits. We showed that blockade of the pIC NMDAr completely prevented conditioned inhibition learning, which to our knowledge is the first evidence indicating any brain region as necessary for acquiring a conditioned fear inhibitor. We conducted a number of experiments, which allow us to conclude that the pIC plays a specific role in learning a conditioned inhibitor, but not in the learning or recall of fear discrimination. These findings

have important implications for understanding the neural regulation of fear and introduce a number of new questions about the functions of IC in cognition. Importantly, other reports fail to find a necessary role of the IC in fear expression (Rosen et al., 1992; Shi and Davis, 1999; although see Casanova et al., 2016 for exception).

The reciprocal connectivity between pIC and the BLA (Shi and Cassell, 1998a, b) positions this structure as a critical intersection for incoming sensory cues to be compared with learned associations. With additional conditioning, an association between the safety signal and the nonoccurrence of shock gains strength because of negative prediction errors generated on no shock trials. That pIC inactivation by muscimol reduced the fear response to the A (Figure 2.11B) suggests that IC may play a role in danger expectation which could be relayed to the amygdala to provide a basis for a prediction error on no shock trials. Although the effect of muscimol on fear recall contrasts with studies that found no critical role of IC in simple fear conditioning, our result is consistent with other findings that over time, a fear stimulus undergoes systems consolidation such that the conditioned response becomes dependent on the IC (Izquierdo et al., 1997). The systems consolidation view also accounts for the discrepancy that inactivation of the IC during recall of conditioned inhibition produced a reduction in fear to the A cue after several days of conditioning, but had no effect on the acquisition or recall of fear discrimination on days 1 or 2 of testing (Figure 2.9).

2.4.4 Relevance to PTSD

The findings of these studies align well with the known neural abnormalities observed in individuals with PTSD. The role of vlOFC in safety signal recall suggest that a consequence of reduced OFC volume and hypoactivity, which are well documented in PTSD (Liberzon and Martis, 2006; Hakamata et al., 2007), is an impairment in switching out of a fearful state even when well-established safety cues are available. In the case of the experiments here, OFC inactivation resulted in a prevailing danger response, despite safety information, and this may explain why impairments in OFC function are commonly observed in PTSD.

Individuals with PTSD also display irregular IC activation, which may interfere with appropriate modulation of fear (Etkin and Wager, 2007; Zhang et al., 2016). Although safety learning has received only limited attention in human neuroimaging studies, Schiller et al. (2008) found increased IC activity during presentation of a danger cue compared to a safe cue. IC activity is also positively correlated with expectations of danger (Phelps et al., 2001) and pain (Ploghaus et al., 1999). While IC does not appear necessary for basic fear discrimination, development of treatments that normalize insular cortex activity may allow individuals with PTSD to better utilize environmental safety cues and modulate fear responding.

2.4.5 Conclusions

Altogether, these mechanistic studies answered important questions about possible nodes in the safety learning circuit. We provide evidence that VH is likely not part of the safety learning mechanism, although it does seem to play a role in acquisition of fear. Conversely, vlOFC appears specifically necessary for the flexible behaviors observed in response to presentations of alternating learned safety and danger cues. More studies on vlOFC will be necessary to narrow in on its role in recall of safety cues. A thorough investigation of IC has allowed us to conclude that pIC, but not aIC or mIC, is necessary for the acquisition of a conditioned inhibitor, as measured by a summation test. Results show that pIC is not necessary for the acquisition or recall of prerequisite fear discrimination. This data may indicate that IC becomes part of the safety circuit when safety signals are well learned, or that the mechanisms that underlie conditioned inhibition of fear are distinct from the circuitry that allows for discrimination between safe and danger cues. The results from these studies greatly informed the hypothesized circuitry underlying fear discrimination presented in Chapter 1, which is further explored in Chapter 4.

CHAPTER 3

Sex Differences In Fear Discrimination Do Not Manifest As Differences In Conditioned Inhibition

The work in this chapter is published in the following manuscript:

Foilb AR, Bals J, Sarlitto MC, Christianson JP (2018) Sex differences in fear discrimination do not manifest as differences in conditioned inhibition. Learn Mem, 25(1): 49-53.

3.1 Introduction

Discrimination between safety and danger is necessary for survival. Incorrect evaluation of a stimulus as safe when it is dangerous could result in harm, while determining a stimulus as dangerous when it is safe results in unnecessary fear and anxiety. Further, when a safety cue is well learned, it can reduce fear in the presence of a danger cue, a learning phenomenon known as conditioned inhibition of fear (Kazama et al. 2013). Overgeneralization of fear-related cues and aberrant conditioned inhibition of fear are seen in individuals with post-traumatic stress disorder (PTSD; Jovanovic et al. 2012; Costanzo et al. 2016; Jenewein et al. 2016). Females are more likely to be diagnosed with PTSD than males (Kilpatrick et al. 2013; Kessler et al. 1995). Safety learning and discrimination are sensitive to hormonal birth control (Lornsdorf et al. 2015) and trauma history in females compared to males (Gamwell et al. 2015). Translational research regarding sex differences in rodent fear discrimination is in its infancy, but has indicated greater discrimination in females compared to males, with later generalization of fear (Day et al. 2016). Biological sex is a significant factor in the expression of fear based psychoses (Shansky 2015) and a better understanding basic behavioral differences in fear discrimination is needed to more fully realize the impact of sex for an individual's health (Shansky and Woolley 2016).

3.2 Methods

Male and female adult Sprague-Dawley rats were obtained from Taconic Biosciences (Hudson, NY) weighing 200-250g upon arrival. Rats were same-sex pair housed in plastic tub cages with free access to food and water, and maintained on a 12h light/dark cycle. Males and females were housed in the same colony room, and were run through behavioral procedures simultaneously. All animals were given 7 days to acclimate to colony housing before behavioral procedures. All experimental procedures were reviewed and approved by the Boston College Institutional Animal Care and Use Committee.

Both conditioning and recall testing occurred in the same, dark context: a 10 x 11 x 6in (L x W x H) chamber made of black plastic with wire mesh lids with a stainless steel grid floor. The chamber was housed within a 15 x 12 x 27in (L x W x H) light and sound-attenuating chamber with a fan for ventilation and background noise (~55dB). We have previously shown the fear discrimination and conditioned inhibition develop equally when tested in the conditioning context or a distinct context. To focus on cue discrimination here, the same context was used for all treatments. Infrared illuminators allowed for video observation through an overhead camera. Scrambled foot shock was delivered via shocker Model H13-15, Coulbourn Instruments. The stimuli were distinct cues of different modalities: a flashing white LED light cue was 264.0 Lux, 20ms on/off and the white noise pip was 10ms duration, 3 Hz interval, 75 dB. Assignment of stimuli as

danger (A) or safe (B) was counterbalanced, and no effect of cue was observed, as reported previously (Chen et al. 2016; Foilb and Christianson 2016; Foilb et al. 2016). Freezing behavior was detected with ANY-Maze computer software (version 4.99, Stoelting, Wood Dale, IL).

Intact male and normally cycling female adult rats were used as follows: n = 24/sex for both fear discrimination acquisition and later fear discrimination recall, n = 12/sex for fear discrimination acquisition and conditioned inhibition, and an additional n = 24/sex that were given fear discrimination acquisition only, for a total N = 120. AX+/BX- discrimination conditioning adapted from Myers and Davis (2004) was used (as in Chapter 2, Chen et al. 2016; Foilb and Christianson 2016; Foilb et al. 2016; Sarlitto et al., 2018). Conditioning trials began with a 5s, 1 kHz (75dB) tone, followed by either A or B for 15s. A trials co-terminated with a 500ms, 1.2mA scrambled foot shock and B trials signaled the absence of shock. Conditioning consisted of 15 presentations of each cue in quasi-random order, so that neither cue occurred more than twice in series with a 90s inter-trial-interval.

Behavioral research has noted sex differences in fear expression, where many female rats exhibit an active response—termed darting—when presented with fear stimuli, as opposed to the passive freezing response typically observed in males (Gruene et al. 2015a). A trained observer screened videos for evidence of darting, but in our experimental conditions darting occurred too infrequently to analyze. Therefore, freezing was used as a behavioral measurement of fear (Fanselow 1980) as previously (Chapter 2, Christianson et al. 2011; Chen et al.

2016,; Foilb and Christianson 2016; Foilb et al. 2016). Freezing data were analyzed by analysis of variance (ANOVA) with sex as a between subjects factor and cue type and trial as within subjects factors with Sidak post hoc contrasts.

3.2.1 Experiment 1: Fear Discrimination

Experiment 1 investigated sex differences in fear discrimination acquisition (n = 60/sex) and recall (n = 24/sex). Fear discrimination recall was tested the day after the initial conditioning session. The test began with 2 min of baseline context exposure, followed by 30s presentations of A and B cues without presentation of shock. Each cue was presented 10 times in pseudo-random order with a 30s inter-trial-interval. To further compare the discrimination abilities of males and females, a discrimination index was calculated as time freezing to A divided by time freezing to B multiplied by 100, so that a discrimination index of 100 indicated equal freezing to A and B cues and lower discrimination scores indicated greater discrimination between cues.

3.2.2 Experiment 2: Conditioned Inhibition of Fear

Experiment 2 investigated sex differences in the acquisition and recall of conditioned inhibition (n = 12/sex). Rats received 5 conditioning sessions as previously (Chapter 2; Foilb et al. 2016). Summation tests were similar to the recall test used in Experiment 3.1 with AB trials added in which A and B were presented in compound. The tests began with 2 min of baseline context, followed

by 30s of each cue—A, B and AB —three times each in randomized order with a 30s inter-trial-interval. To further observe differences in conditioned inhibition, a summation index was calculated as time freezing to AB divided by time freezing to A multiplied by 100.

3.3 Results

3.3.1 Experiment 1: Fear Discrimination

Females displayed discrimination earlier in conditioning than males (Figure 3.1A). A 3-way ANOVA of sex by cue by trial block (5 blocks of 3 cue trials) revealed a main effect of trial block, $F(4, 472) = 40.579$, $p < 0.001$, and interactions of trial by sex, $F(4, 472) = 3.491$, $p = 0.008$, and cue by trial, $F(4, 472) = 45.474$, $p < 0.001$, but no significant interaction of trial block by cue by sex interaction, $F(4, 473) = 1.048$, $p = 0.382$. Post hoc comparisons showed that females significantly discriminated between the A and B in the first trial block ($p = 0.007$), while males failed to show to this discrimination until the second trial block ($p = 0.769$ on trial block 1; $p = 0.004$ on trial block 2). Females also displayed significantly reduced freezing to A compared to males on trial blocks 3, 4, and 5 ($ps = 0.039$, 0.018, 0.003, respectively) as well as less freezing to B compared to males on all trial blocks ($p = 0.048$ on the first trial block, $ps < 0.001$ on trial blocks 2, 3, 4, and 5).

Freezing to each cue during conditioning is summarized in Figure 3.1B. ANOVA revealed main effects of sex, $F(1, 118) = 23.42$, $p < 0.0001$, and cue, F

(1, 118) = 312.3, $p < 0.0001$, and a cue by sex interaction, F (1, 118) = 22.15, $p < 0.0001$. Post hoc analyses showed that both sexes significantly discriminated between A and B, ps < 0.0001, but that females displayed significantly less freezing to B, $p < 0.0001$, compared to males. A discrimination index was calculated (freezing to B / freezing to A x 100) as a measure of animals' ability to discriminate between the A and B (Figure 3.1C). Males and females showed significantly different discrimination indices, t (118) = 6.343, $p < 0.0001$, with females showing more discrimination compared to males.

Acquisition of fear discrimination was established in 3 cohorts of animals. Cohort 1 are animals that received fear discrimination recall (n = 24; Figure 3.2), cohort 2 received conditioned inhibition recall testing (n = 12; Figure 3.3) and cohort 3 were animals run through the same conditioning procedures with all the same procedures, but do not have recall data included here (n = 24). The sex difference observed in fear discrimination acquisition was consistent across all cohorts of animals (Figure 3.1D). Analysis of discrimination indices from each cohort found a main effect of sex, as expected, F (1, 114) = 37.48, $p < 0.0001$. Importantly, there was no main effect of cohort, F (2, 114) = 2.325, $p = 0.1024$, and no sex by cohort interaction, F (2, 114) = 0.072, $p = 0.93$. Post hoc analyses revealed that males and females had significantly different discrimination indices in cohort 1, $p < 0.0001$, cohort 2, $p = 0.01$, and cohort 3, $p = 0.0008$. Additional analyses by acquisition cohort are provided in the published manuscript (Foilb et al., 2018).

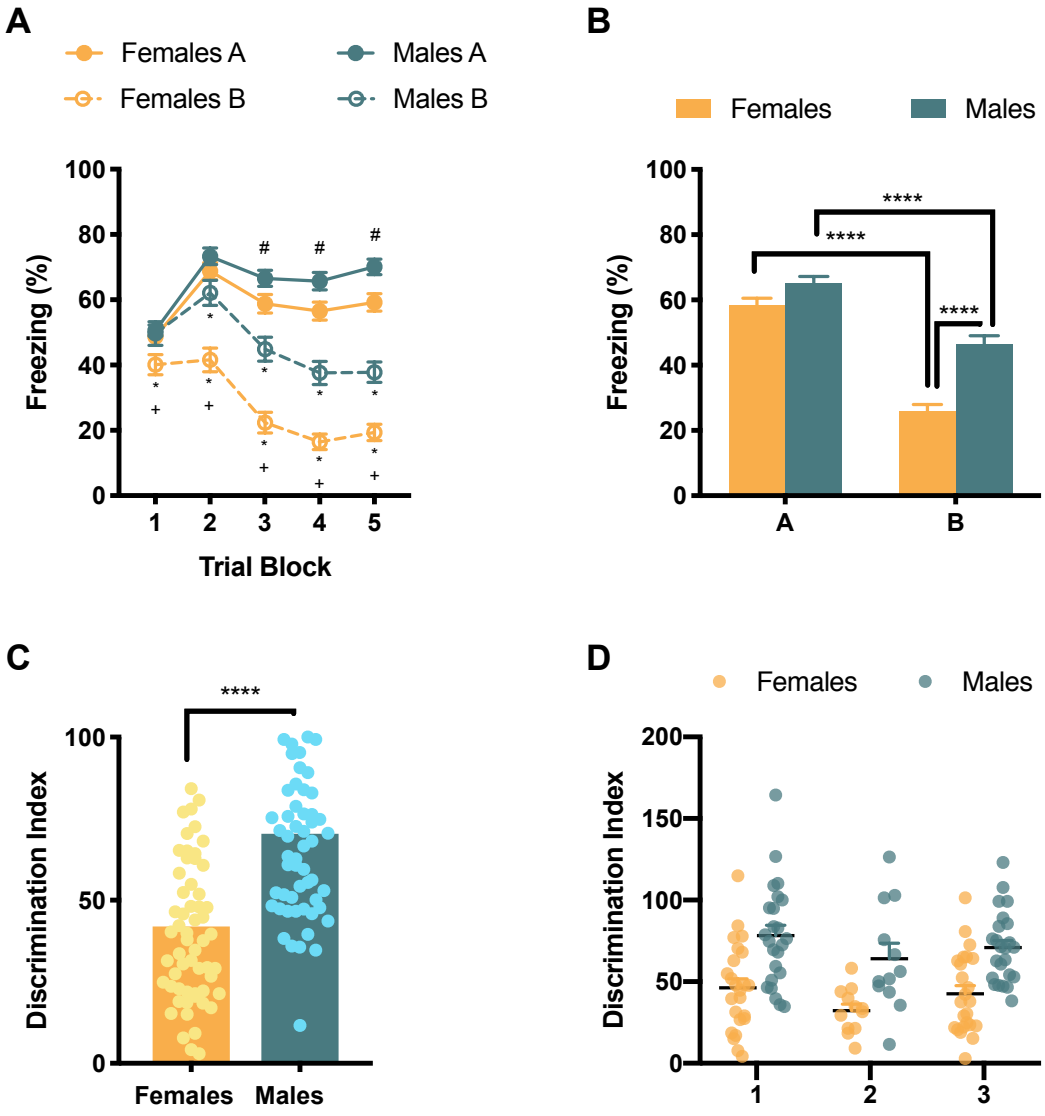

Figure 3.1. Sex differences in fear expression and discrimination during AX+/BX- conditioning. (A) Freezing averages (±SEM) to A and B across blocks of three trials in conditioning. Females significantly discriminated between A and B within the first trial block of each cue ($p = 0.001$), while males did not make this discrimination until the second trial blocks ($p = 0.002$). Females also displayed less freezing to B than males on all trial blocks ($ps < 0.05$). *significant difference between A and B within sex, $^{+}$significant difference between males and females to B, and $^{\#}$significant difference between sexes on A. **(B)** Mean (+SEM) freezing to A and B cues during conditioning. Both males and females significantly discriminated between A and B ($ps < 0.0001$), but females displayed significantly less average freezing to B ($p < 0.0001$) compared to the males. **(C)** Mean (and individual replicates) discrimination indices (time freezing to B / time freezing to A x 100) during conditioning. **(D)** Mean (±SEM, individual replicates) discrimination indices during conditioning of females and males in each of the 3 cohorts. Females displayed significantly lower discrimination indices compared to males, indicating more discrimination, in all 3 cohorts, $ps < 0.05$. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$

Sex differences were also evident in fear discrimination recall tests. An ANOVA of trials across the test (Figure 3.2A), revealed a main effect of trial, $F(9, 414) = 14.856$, $p < 0.001$, and cue by trial interaction, $F(9, 414) = 3.333$, $p = 0.001$, but no significant interactions of sex by trial, $F(9, 414) = 0.421$, $p = 0.924$, or cue by trial by sex, $F(9, 414) = 1.668$, $p = 0.095$. Females significantly discriminated between A and B on the first trial ($p < 0.001$), and continued this level of discrimination throughout the test. Males also displayed immediate discrimination on trial 1 ($p = 0.004$) and matched the discrimination level of the females throughout the remainder of the test ($p < 0.001$). While females' response to B was stable across trials, males reduced freezing on B presentations 7, 8 and 9 compared to trial 1 (ps = 0.004, 0.003, and < 0.001, respectively), as well as on trial 10 compared to trials 2 ($p = 0.048$) and 3 ($p = 0.037$), and on A presentations 8, 9 and 10 compared to presentation 4 (ps = 0.034, 0.03 and 0.007, respectively). Males and females significantly differed in their response to A on trials 1 ($p = 0.019$), 6 ($p = 0.038$), 7 ($p = 0.019$), 8 ($p = 0.049$), and 9 ($p = 0.020$) and showed even more differential responding to B, with significantly different freezing on trials 1 to 6 and trial 8 (ps < 0.032).

Figure 3.2B shows average freezing to each cue during recall. ANOVA revealed main of effects of cue, $F(2, 92) = 93.01$, $p < 0.0001$, sex, $F(1, 46) = 40.33$, $p < 0.0001$, and cue by sex interaction, $F(2, 92) = 12.74$, $p < 0.0001$. Post hoc analyses showed that females continued to freeze significantly less than males to B ($p = 0.0003$), as well as to the baseline context at the start of the test

($p < 0.0001$). Females froze significantly less to context than all other cues (ps < 0.01) and froze significantly less to B than to A ($p < 0.0001$). Males also froze significantly less to B and baseline context compared to A (ps < 0.0001), but freezing to B and baseline context did not significantly differ ($p = 0.641$). This sex by cue interaction is summarized by significant difference in discrimination index (Figure 3.2C, $t(46) = 3.63$, $p = 0.0007$).

To determine if the sex differences in cue response were not simply artifacts of differential baseline fear, we subtracted baseline freezing from average freezing to each cue (Figure 3.2D). Here we found main effects of cue, $F(1, 46) = 248.6$, $p < 0.0001$, sex, $F(1, 46) = 14.05$, $p = 0.0005$, and a cue by sex interaction, $F(1, 46) = 8.692$, $p = 0.005$. These differences indicate that the sex difference in discrimination is not only due to the sex difference in baseline fear, but that there is a general sex difference in discrimination between the CS and the conditioning context.

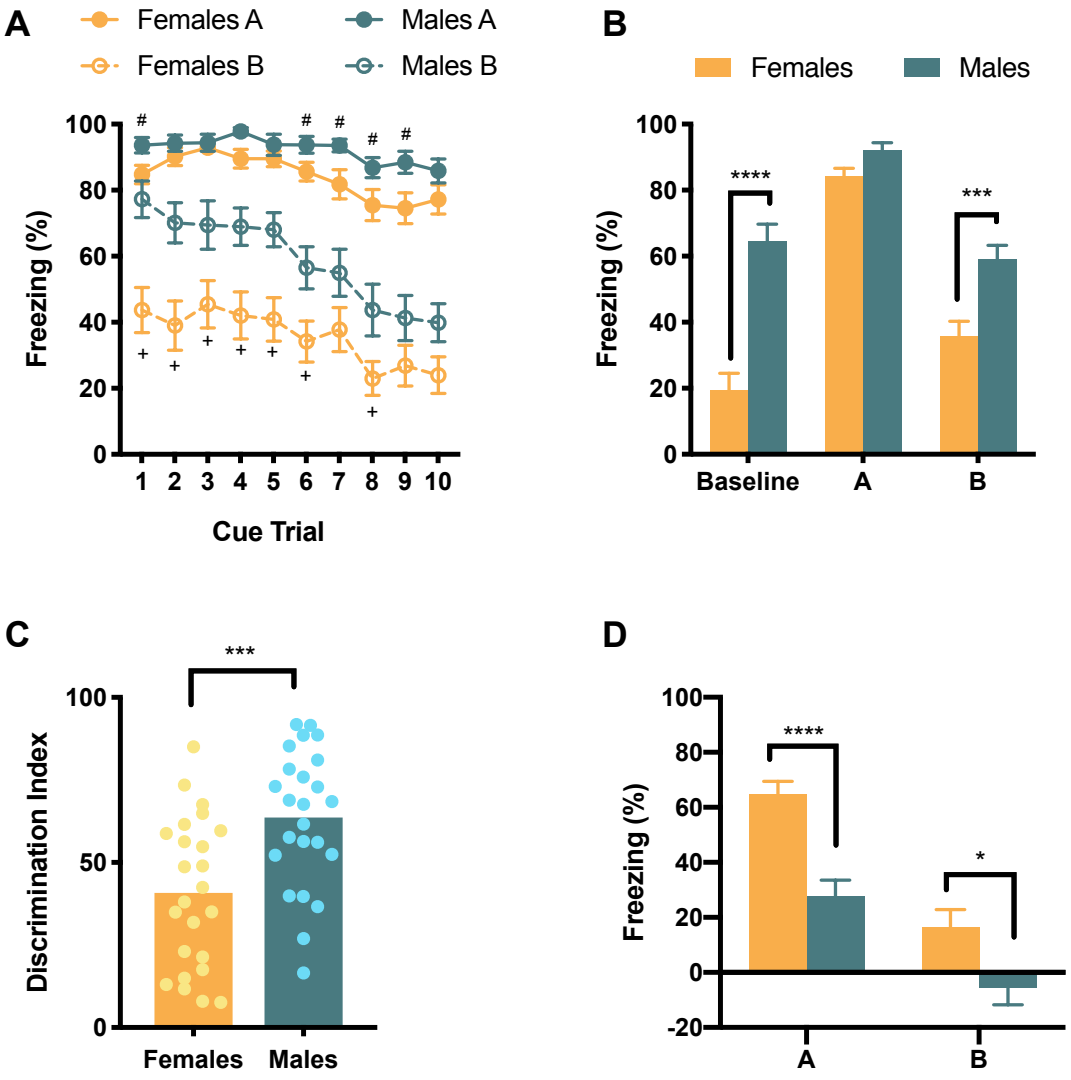

Figure 3.2. Sex differences in fear discrimination recall. (A) Mean freezing (±SEM) to A and B across each of the 10 cue presentations in testing. Both sexes significantly discriminated between A and B on all trials (significance not marked on graph). Females displayed significantly less freezing to B compared to males at the start of the test, while males reduced freezing to B throughout the test. $^{+}$significant difference between sexes to B ($ps < 0.05$), and $^{\#}$significant difference between sexes to A ($ps < 0.05$). **(B)** Average freezing (+SEM) to baseline context exposure, A and B during the recall test. Females displayed significantly less freezing to the baseline context and B compared to males. There was no sex difference in freezing to A. **(C)** Mean (individual replicates) discrimination indices in recall testing. Females have a significantly lower discrimination index compared to males. **(D)** Average freezing (+SEM) to A and B with baseline freezing subtracted. When comparing A to B freezing, males and females showed significant discrimination (A vs. B $ps < 0.0001$), while females showed greater freezing to the baseline subtracted A and baseline subtracted B compared to males ($p < 0.0001$ and $p < 0.05$, respectively).* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$

3.3.2 Experiment 3.2: Conditioned Inhibition of Fear

Conditioned inhibition of fear was assessed in summation tests. On test 1, analysis on average freezing to each cue (Figure 3.3B) in a 4 (cue) by 2 (sex) ANOVA revealed a significant effect of cue $F(3, 66) = 30.06$, $p < 0.0001$, with significantly reduced freezing to B compared to the A and AB (ps < 0.0001) and significantly increased freezing to A and AB compared to baseline context (ps < 0.0001). Animals did not discriminate between A and AB in this test ($p = 0.33$). There was no main effect of sex, $F(1, 22) = 0.1599$, $p = 0.6936$, or cue by sex interaction, $F(3, 66) = 0.8828$, $p = 0.4547$. The lack of sex difference in A/B discrimination is likely the result of differences in cue presentation in the summation test compared to the recall test. Specifically, the first presentation of the B cue in the summation test is as a compound with A. In test 5 (Figure 3.3C), the same analysis revealed a significant main effect of cue, $F(3, 66) = 35.72$, $p < 0.0001$, and a significant cue by sex interaction, $F(3, 66) = 3.15$, $p = 0.0307$, but no main effect of sex, $F(1, 22) = 0.1521$, $p = 0.7003$. Animals displayed greater fear to A compared to the baseline context, B and AB (ps < 0.0001). Animals also froze less to the B compared to the AB and baseline context (ps < 0.05).

Comparing discrimination indices across all 5 summation tests (Figure 3.3D) with a 5 (test day) by 2 (sex) ANOVA revealed a main effect of test day, $F(4, 88) = 5.603$, $p = 0.0005$, but no main effect of sex, $F(1, 22) = 1.328$, $p = 0.2615$, and no test day by sex interaction, $F(4, 88) = 0.9538$, $p = 0.4371$. Discrimination between A and B was significantly improved on tests 3, 4, and 5

compared to test 1 (ps < 0.05). A summation index was calculated as freezing to AB / freezing to A x 100 as a measurement of conditioned inhibition (Figure 3.3E). ANOVA revealed a main effect of test day, F (4, 88) = 14.9, $p < 0.0001$, but no main effect of sex, F (1, 22) = 1.949, $p = 0.1767$, and no test day by sex interaction, F (4, 88) = 1.751, $p = 0.1460$. Rats showed greater inhibition on tests 4 and 5 compared to tests 1, 2, and 3 (ps < 0.01).

The difference in baseline context freezing that was observed in Experiment 1 was present in this smaller sample as a trend, but did not reach significance. This may reflect an effect of different estrous status on the test day between experiments, or sampling error and the intrinsic variability in this dependent measure, which is assessed in Figure 3.4. A two-way ANOVA found a main effect of sex, F (1, 68) = 16.38, $p < 0.001$, and an experiment by sex interaction, F (1, 68) = 5.028, $p = 0.0282$, but no main effect of experiment, F (1, 68) = 0.1098, $p = 0.7414$. Although freezing levels changed in males and females between Experiments 1 and 2, the mean values did not differ from each other (p = 0.141 and $p = 0.3294$, respectively; Figure 3.4A). Analysis of baseline freezing by sex in test 1 of Experiments 1 and 2 combined, revealed that males and females displayed significantly different baseline freezing across the two tests, t (70) = 4.975, $p < 0.0001$. Together these analyses show that despite differences in significance at baseline freezing in Experiments 1 and 2, the variability of freezing and overall pattern of reduced baseline freezing in females was consistent throughout both recall experiments.

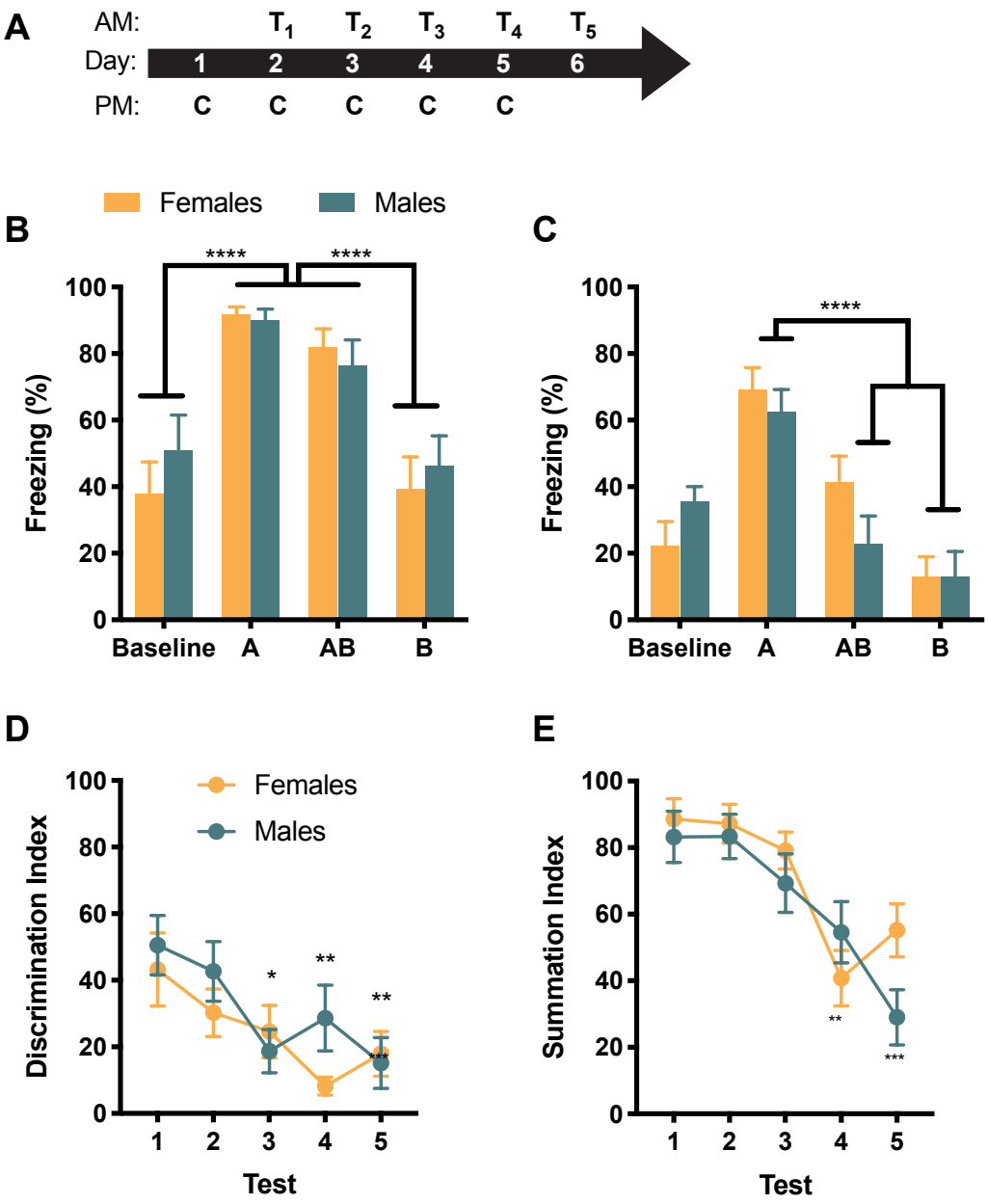

Figure 3.3. Sex differences were not evident in conditioned inhibition of fear. (A) Timeline of experiment where recall testing (T) occurred the morning after each conditioning session (C). **(B)** Average freezing (+SEM) to baseline context exposure, A, AB and B during summation test 1. Animals significantly discriminated between A and B ($p < 0.0001$), but did not discriminate between A and AB ($p = 0.33$). **(C)** Average freezing (+SEM) to baseline context exposure, A, AB and B during recall test 5 on day 6. Animals significantly discriminated between all cues, with reduced freezing to AB and B compared to A ($ps < 0.0001$). **(D)** Discrimination indices (±SEM) across each of the 5 tests. While there was no sex difference on any test day, discrimination improved on tests 3 ($p < 0.05$), 4 and 5 ($ps < 0.01$) compared to tests 1 and 2. **(E)** Summation indices (±SEM) across each of the 5 tests. There were no sex difference in summation, but improved inhibitory summation on tests 4 and 5 compared to tests 1, 2 and 3 ($ps < 0.01$). * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$

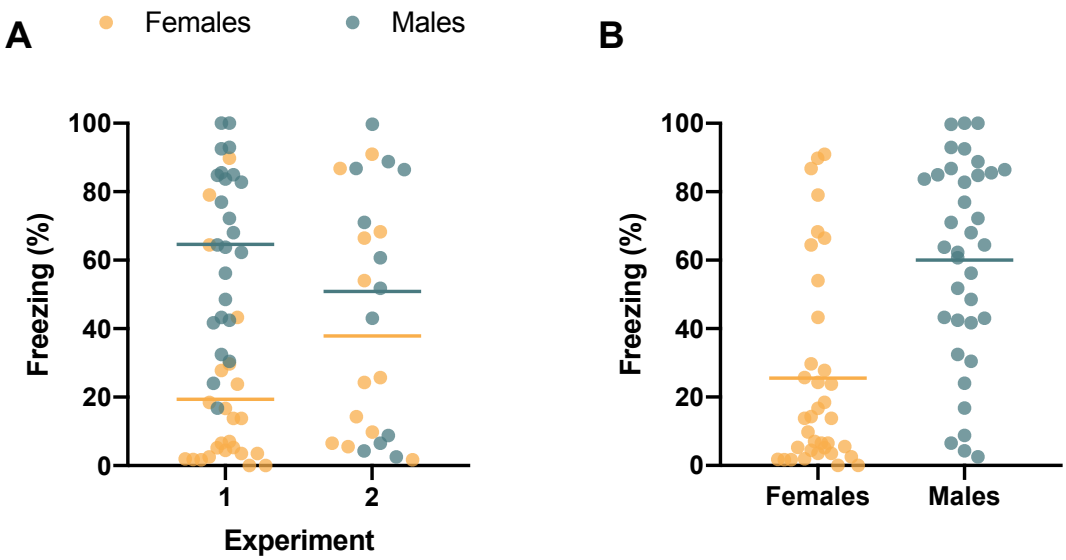

Figure 3.4. **Baseline freezing in Experiments 1 and 2. (A)** Mean (individual replicates) baseline freezing in recall tests of Experiment 1 and the first test of Experiment 2. While baseline freezing levels changed between Experiments 1 and 2, the mean values for each sex did not significantly differ from each other. **(B)** Mean (individual replicates) baseline freezing by sex in test 1 of Experiments 1 and 2 combined. Males and females displayed significantly different baseline freezing across the two tests, t (70) = 4.975, $p < 0.0001$.

3.4 Discussion

We observed a marked difference in fear discrimination between male and female rats. Differential freezing to the danger A cue and safe B cue was greater in females compared to males during the initial conditioning and in a recall test one day later. Furthermore, females exhibited greater initial discrimination between the context and the discrete cues. These results are consistent with recent findings (Day et al. 2016), and we were able to replicate our findings in

discrimination acquisition, as presented in Chapter 4. Although the sex differences we observe are in contrast to studies on fear discrimination in humans where females show less discrimination than males (Gamwell et al. 2015; Lornsdorf et al. 2015). Females also displayed lower contextual freezing, consistent with several prior reports in rodents (Pryce et al. 1999; Daviu et al. 2014; Pettersson et al. 2016; although see Keiser et al. 2017 for exception). Attempts to uncover the mechanisms underlying reduced contextual fear in females have mostly resulted in evidence to the contrary. Exploration in the hypothalamic-pituitary-adrenal (HPA) axis of stress response, found that females have a greater hormonal stress response during fear conditioning compared to males, despite the reduced fear expression and faster extinction rates seen in females (Daviu et al., 2014). Interestingly, when this difference in contextual freezing was subtracted from freezing to each cue in the experiments presented here, there was a larger difference in freezing to A, whereas the main difference in the uncorrected data was primarily in freezing to B. This indicates that the sex difference may be in discrimination *per se*, rather than freezing to a particular cue.

The discrimination test requires that rats flexibly transition between fear and safe states, which may favor the inherently more active female rats (Gruene et al. 2015a, b). However, after several days of conditioning, females may accrue more fear to the safe cue (Day et al. 2016) which could manifest as a transition from more active behavior early in conditioning to more passive, i.e. male-like,

with additional conditioning and stress (Foilb and Christianson 2016) and account for the lack sex difference in conditioned inhibition summation tests. However, this outcome contrasts some of the results of Day and colleagues (2016) in which after repeated discrimination conditioning the safety cue failed to pass a retardation test of conditioned inhibition. Whether these different empirical results are a consequence of procedural differences or of different neural mechanism underlying summation and retardation phenomena remains unknown.

The sex difference in discrimination indicates that there may be sex differences in the neural circuitry underlying discrimination learning. Interactions between the amygdala and prefrontal cortex are critical for danger/safety discrimination (Likhtik et al. 2014) and sexual dimorphisms observed in humans include sex differences in amygdala anatomy (Ruigrok et al. 2014), amygdala response to negative or stressful emotions (Stevens and Hamann 2012; Kogler et al. 2015), and amygdala functional connectivity (Lopez-Larson et al. 2011; Engman et al. 2016). Sex differences have been found in the rodent medial prefrontal cortex (Baran et al. 2010; Fenton et al. 2014; Fenton et al. 2016), a brain region critical to fear expression, extinction and discrimination (Sotres-Bayon and Quirk 2010; Milad et al. 2014; Sangha et al. 2014). Similarly, basolateral amygdala projecting neurons of the infralimbic region of the medial prefrontal cortex appear to differently mediate fear expression in males and females (Gruene et al. 2015b).

In contrast to a circuit-based substrate for sex differences in fear discrimination prior work regarding gonadal hormones and fear discrimination would predict that females in the current test would perform more poorly than males. The addition of estrogen did not alter discrimination in gonadectomized males, but prevented conditioned inhibition of fear in ovariectomized females (Toufexis et al., 2007). It has also been reported that estradiol replacement to ovariectomized rats leads to increased generalization of contextual fear (Lynch et al., 2013). Estrogen receptor agonists also lead to disruption of fear in both males and females, with increased overall fear when the estrogen receptor α, a receptor implicated in fear and anxiety, was agonized, but not when the β receptor, known to be anxiolytic, was agonized (Toufexis et al., 2007; Morgan & Pfaff, 2001; Walf, Rhodes, & Frye, 2004; Walf & Frye, 2005). Yet preliminary estrous testing of the females here showed no trend between cycle phase and ability to acquire or recall the safety signal, so cycle data collection was stopped in order to maintain equal treatment of sexes. Human literature also contrasts the existing female rodent data, and a report in women found estrogen levels were positively correlated with fear inhibition (Glover et al. 2013). This may simply suggest that hormone replacement studies do not completely inform the role of gonadal hormones on fear in regularly cycling, intact females.

That sex differences in safety learning do not persist in conditioned inhibition of fear indicates that conditioned inhibition of fear occurs in a neural circuit that is distinguishable from the fear discrimination circuitry. Indeed, the

discrimination task employed here requires that rats rapidly transition between fear and safe states which suggests inputs that alter the output of the amygdala could be recruited in sex-specific ways. Sex differences in neuronal activation are likely to be most robust during the early phase of safety learning and future studies seeking to understand how the biological variable of sex shapes the function the fear circuitry to better inform and individualize treatments for fear based psychiatric diseases.

CHAPTER 4

Neural Correlates of Safety Learning

4.1 Introduction

Inappropriate expression of fear or anxiety, or fear or anxiety under conditions that would not normally elicit fear in a healthy individual, is a distinguishing characteristic of post-traumatic stress disorder (PTSD; Rauch et al., 2000). One case in which fear may be expressed inappropriately is in the presence of information indicating safety rather than danger. Accurate discrimination between safety and danger allows an individual to respond flexibly when threat is most likely to occur. In a laboratory setting with rodents, danger is established by presenting a neutral stimulus (the conditioned stimulus, or CS) paired with a mild electric shock (the unconditioned stimulus, or US). When the CS is later presented, it elicits fear, often observed as behavioral freezing in laboratory rats (Fanselow, 1980). Safety signals acquire value through associative learning when there is a non-zero probability of an aversive stimulus. Safety learning occurs more gradually than danger learning, requiring multiple presentations of both danger cues (A) followed by an aversive US and safety cues (B) indicating the absence of shock.

Fear discrimination has been studied in laboratories across species, where humans, monkeys and rodents are all able to accurately discriminate between safety and danger (Jovanovic et al., 2005; Winslow et al., 2008; Myers and Davis, 2004). Yet difficulty utilizing learned safety signals has been observed in a number of clinical PTSD studies (Jovanovic et al., 2012; Costanzo et al., 2016; Jenewein et al., 2016). Despite the increased prevalence of PTSD in

females compared to males, much of the research on fear discrimination has focused on males (Kessler et al., 1995; Kilpatrick et al., 2013). In humans, fear discrimination is more sensitive to trauma history in females compared with males (Gamwell et al., 2015). Fear discrimination in females is also altered by hormonal birth control, indicating that sex hormones may play a role in sex differences in discrimination (Lonsdorf et al., 2015). Translational research regarding sex differences in rodent fear discrimination is relatively limited, but a burgeoning area of exploration. Reports, including work from our own lab, indicate greater discrimination in females compared with males, with female rodents able to modulate fear to a safety signal more rapidly than males (Day et al., 2016; Foilb et al., 2018).

Compared to related areas of research, such as fear conditioning and extinction learning, there is limited work on the neural circuitry of cued fear discrimination. A better understanding of the neural mechanisms that underlie safety learning will provide critical information relevant to disorders of abnormal fear modulation. Based on existing information about the neural structures involved in fear learning and fear discrimination, there is a clear set of brain regions likely to be involved in safety learning; these structures are the focus of this work. Prefrontal regions known to be involved in the modulation of fear, including the prelimbic and infralimbic regions of the prefrontal cortex (PL and IL, respectively; Sotres-Bayon and Quirk, 2010; Sierra-Mercado et al., 2011) and the orbitofrontal cortex (OFC; Sarlitto et al., 2018), are likely components of fear

discrimination circuitry. IC is known to be a region of emotion learning and salience detection (for review, see Gogolla, 2017); processes closely related to the acquisition of safety signals. IC has also been found necessary for the stress mitigating effects of safety signals (Christianson et al., 2008, 2011). It is important to look at regions known to be involved in fear learning and expression, such as the basolateral amygdala (BLA; Quirk et al., 1995; Rogan et al., 1997; LeDoux, 2000; LeDoux, 2014), the bed nucleus of the stria terminalis (BNST; Campeau et al., 1997; Walker and Davis, 2008; Davis et al., 2010), and the central amygdala (CeA; LeDoux et al., 1988; Swanson and Petrovich, 1998; Maren, 2001). These regions related to fear learning and expression are of particular interest since there is also evidence for the involvement of these regions, BLA and BNST in particular, specifically in safety learning (Genud-Gabai et al., 2013; Sangha et al., 2013; Ostroff et al., 2010, 2012; Campeau et al., 1997; Christianson et al., 2011). The evidence that each of these regions is involved in safety learning and the interconnection of these regions in ways that are likely meaningful to the process underlying this behavior, are elaborated on in the following section. Together, these brain regions may be key nodes in a brain wide safety learning circuit, which is the basis of this work (Figure 4.1).

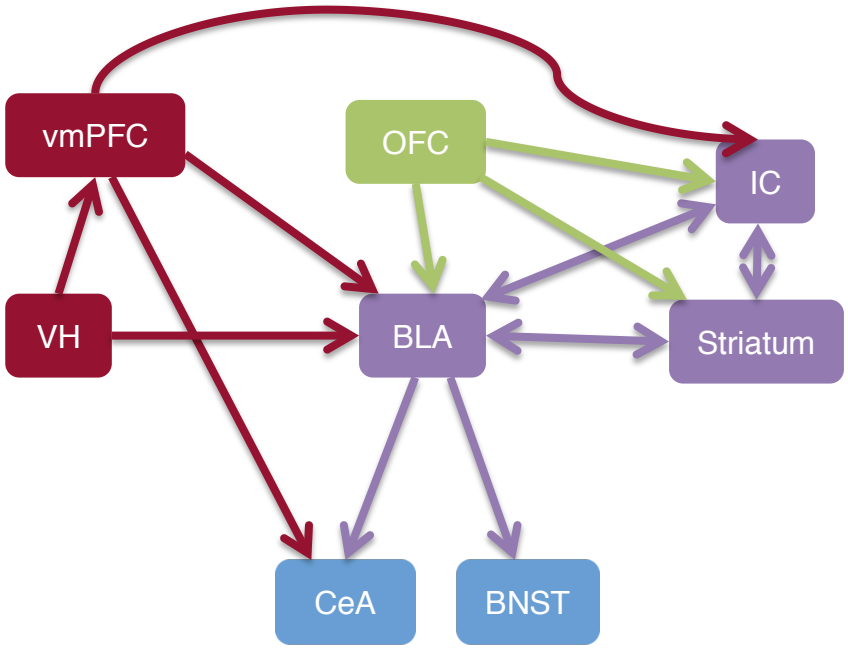

Figure 4.1 A hypothetical circuit for the processing of safety information. This circuit is the basis of the work presented here. Red regions and arrows indicate a site of danger processing and the projection of danger information. Green regions and arrow indicate regions and transferring of safety information. Purple colored regions and arrows indicate regions that display altered patterns of responding due to the reception of safety information. Blue regions indicate regions that project to the regions, which ultimately lead to behavioral outputs.

4.1.1 Fear Modulation: Prefrontal Cortex

The ventral medial prefrontal cortex (vmPFC) plays a critical role in the modulation of fear, with compelling support for fear promoting and fear inhibiting roles of the PL (Sotres-Bayon and Quirk, 2010) and IL (Sierra-Mercado et al., 2011), respectively. The involvement of vmPFC in fear is mediated through projections from PL and IL to specific amygdala subregions, as well as inputs from amygdala, thalamus and hippocampus (Sotres-Bayon and Quirk, 2010). Structural differences in the vmPFC are also reported in PTSD populations, indicating that this could be a structure critical for appropriate modulation of fear

(Corbo et al., 2005; Etkin and Wager, 2007; Hughes and Shin, 2011). Sangha and colleagues (2014) looked individually at PL and IL in a fear discrimination paradigm. Inactivation of PL led to a reduction of freezing to the danger cue A, but did not alter freezing to B or AB cues compared to vehicle animals. Consistently, Likhtik and colleagues (Likhtik et al., 2014) observed a strong correlation between PL and BLA synchrony and behavioral discrimination during a differential inhibition task in mice. Inactivation of the IL before recall testing resulted in reduced freezing to A, disrupting discrimination between the A and AB cues (Sangha et al., 2014). Human evidence also indicates a role of vmPFC in fear discrimination; with increased vmPFC activation to safety cues compared to danger cues (Schiller et al., 2008).

Our lab looked at the role of ventrolateral OFC (vlOFC) in fear discrimination since the region has been implicated in value-based decision-making (Sul et al., 2010), as well as in switching between cognitive tasks (Wilson et al., 2014). OFC is also well connected for involvement in fear modulation—receiving sensory and amygdala inputs and sending projections to IC, amygdala, and striatum (Ongür and Price, 2000; Price, 2007). Based on these functions and connectivity, we hypothesized that vlOFC would be recruited during the changes in behavioral freezing that occur in fear discrimination. Temporary inactivation of the vlOFC with muscimol before a discrimination recall test impaired discrimination resulting in greater fear to the safety cue B. Inactivation of vlOFC during acquisition of fear discrimination did not alter discrimination behavior

during conditioning or at later recall (Sarlitto, Foilb and Christianson, 2018). Despite null results of vlOFC inhibition during acquisition, it is important to note that mechanistic experiments on vlOFC were only performed in males. Since females display greater discrimination than males during conditioning, vlOFC may be a region of differential activation that contributes to faster discrimination learning in females, as it seems to be involved in the switching between appropriate cue responses. Others have also found that lesions of lateral OFC results in generalization of fear in a discrimination learning paradigm (Ray et al., 2018). Together, the existing data indicate that the vmPFC and OFC contribute to different aspects of recall of both danger and safety signals.

4.1.2 Danger/Safety Integration: Insular Cortex

Research from our lab found that posterior IC (pIC) was specifically involved in the acquisition of condition fear inhibition, as measured by a summation test (Foilb et al., 2016). IC may play a role in fear discrimination, as it has access to somatosensory information and is known to be involved in convergent responses to multisensory information (Rodgers et al., 2008). In terms of connectivity, pIC receives projections from vmPFC, and is bidirectionally connected to BLA and CeA, as well as to anterior (aIC) and medial (mIC) regions of IC, which also contain dense connectivity to amygdala (Shi and Cassell, 1998a, b). While aIC and mIC were found unecessary for acquisition of

conditioned inhibition of fear in our previous work, these regions have not been explored in fear discrimination.

4.1.3 Fear Learning and Expression: The Extended Amygdala

The BLA is a critical structure for fear. It is the site of neuroplasticity for fear learning, and is necessary for expression of conditioned fear (Quirk et al., 1995; Rogan et al., 1997; LeDoux, 2000). Many manipulations that prevent BLA excitability or plasticity—including inhibitory drugs, lesions, and optogenetic silencing—all interfere with the learning and later expression of conditioned fear (Maren et al., 1996; Cousens and Otto, 1998; Lalumiere, 2014). Therefore, safety signals might also utilize the BLA for both learning and recall. Many studies have in fact found evidence that safety signals impact neuronal responding in the BLA, in single unit recordings (Genud-Gabai et al., 2013; Sangha et al., 2013), as well as spine morphology and synapse size and strength (Ostroff et al., 2010, 2012). The principle outputs of the BLA that initiate and maintain fear responses are the CeA and the BNST. The CeA receives sensory and visceral information from the BLA and projects to the hypothalamus and brainstem areas responsible for the fear response (LeDoux et al., 1988; Swanson and Petrovich, 1998; Maren, 2001). Redundancy in the output circuitry allows for expression of fear in the absence of CeA functioning via BLA projections to BNST (Campeau et al., 1997; Walker and Davis, 2008; Davis et al., 2010). Quantification of neural activation after feature negative learning, where a safe cue precedes the danger cue on non-reinforced

trials, found that safety trials led to activation in the BNST (Campeau et al., 1997). Conversely, in a backwards conditioning paradigm, where the safety cue signals the end of shock, leads to reduced activation of the BNST, without effect on the CeA (Christianson et al., 2011). These differences are most likely due to the different procedures used, as feature negative and backwards conditioning may engage different neural circuitries. BNST efferents are very similar to those of the CeA and are involved in the sustained expression of fear. BNST also receives dense projections from CeA and IL (Hurley et al., 1991; Dong et al., 2001; Walker and Davis, 2008). Safety signals, therefore, likely alter the expression of fear through a circuit involving the BNST. Nonetheless, looking at CeA and BNST will lead to a better understanding of how safety signals are processed.

4.1.4 Sex Differences

The existing work on the neural mechanisms underlying acquisition and retrieval of safety cues has been done entirely in males. With the behavioral evidence that males and females likely differ in fear discrimination paradigms, it is imperative that mechanistic differences underlying this behavior are explored across both sexes. Biological sex is a significant factor in fear-based psychoses (Shansky, 2015; Shansky and Woolley, 2016). Understanding the neurological differences that underlie sex differences in fear discrimination will allow for a more comprehensive view of how biological sex impacts an individual's health.

Many brain regions of interest are known to be sexually dimorphic. Stress and sex-related hormones impact PFC in sex-specific ways (Farrell et al., 2015, 2016). Sex differences in both vmPFC and amygdala have been shown to play a role in fear learning and extinction (Zeidan et al., 2011; Gruene et al., 2015). BNST is highly sexually dimorphic with larger volume in males in both rodents and humans, which has been hypothesized to be related to consistently observed sex differences in contextual fear (for review, see Goode and Maren, 2017). These structures are also potential mediators of the behavioral sex difference observed in fear discrimination.

Further investigation into sex differences in safety learning may explain why more women than men are diagnosed with PTSD, and address the inability of individuals with PTSD to properly inhibit fear responses – particularly if the mechanisms in females differ from those in males (Jovanovic and Norrholm, 2011; McLean et al., 2011; Lebron-Milad et al., 2012). This work aims to more thoroughly examine sex differences in neural mechanisms underlying safety learning; steps critical to progress in the treatment of PTSD and other anxiety disorders with impairments in fear modulation.

4.1.5 Goal of this Work

Here we explore the brain regions involved in acquisition of fear discrimination in males and females using Fos as a marker of neural activation. Quantifying Fos, the protein product of immediate early gene *c-fos,* which is

induced in response to a large range of stimuli, is widely used as a proximate measure of a neuron's recent activity (Dragunow and Faull, 1989; McReynolds et al., 2018). Since the neural basis of fear discrimination is largely unknown, we compare fear discrimination neural activation patterns to animals that are not given a safety cue (Fear Only group) and to animals that are presented the conditioning stimuli but are never shocked (Control group). This approach allows for observation of brain regions that may be activated specifically due to fear discrimination compared to fear learning without safety information or the sensory experience of cue exposure with no fear-related learning. This approach makes this study particularly impactful since it takes into consideration potential sex differences in neural activity underlying the acquisition of fear discrimination, as well as activity in brain regions that distinguish fear discrimination learning from fear learning or no learning.

We analyzed Fos in several ways to 1) test if fear discrimination conditioning differentially activates regions compared to fear conditioning and controls, 2) test for potential sex differences in Fos levels, 3) to correlate regional Fos with discrimination ability, and 4) to begin to characterize how fear discrimination alters the functional connectivity within the neural regions of interest.

4.2 Materials And Methods

4.2.1 Animals

Rats were obtained from Taconic Biosciences (Hudson, NY). Males and females were weight-matched, arriving at 200-250g, and housed in same-sex pairs in plastic tube cages with free access to food and water. Male and female pairs were housed in the same colony room, where they were kept on a 12 hour light/dark cycle. All animals had 7 days to acclimate to colony housing before any experimental procedures took place. Each of 3 experimental groups included 8 male and 8 female rats. All experimental procedures were reviewed and approved by the Boston College Institutional Animal Care and Use Committee.

4.2.2 Apparatus

All behavioral conditions were performed in the same context: a 15 x 12 x 27in (L x W x H) light and sound-attenuating chamber with a fan for ventilation and background noise (~55dB) housed a 10 x 11 x 6in (L x W x H) chamber made of black plastic with wire mesh lids with a stainless steel grid floor. A 1.2 mA, 0.5s scrambled foot shock was delivered via shocker Model H13-15, Coulbourn Instruments. Digital video cameras (Model VX-5000, Microsoft, Redmond, VA) were used to record behavior, with infrared blocking filters removed. Infrared LEDs illuminated the chambers, allowing for video observation of freezing as a behavioral measure of fear, and detected with ANY-Maze

computer software (version 4.99, Stoelting, Wood Dale, IL). Stimuli were delivered through a flashing white LED light, 264.0 Lux, 20ms on/off, and a white noise pip was 10ms duration, 3 Hz interval, 75 dB. Assignment of stimuli as danger (A) or safe (B) was counterbalanced, and no effect of cue was observed, as reported previously (Foilb et al., 2018; Chen, Foilb and Christianson, 2016; Foilb and Christianson 2016; Foilb et al. 2016; Sarlitto, Foilb and Christianson, 2018).

4.2.3 Behavioral Conditions

Conditioning for the Discrimination group was adapted from (Myers and Davis, 2004) and used previously (Chen et al., 2016; Foilb and Christianson, 2016; Foilb et al., 2016, 2018; Sarlitto et al., 2018), AX+/BX- fear discrimination conditioning consisted of 15 presentations each of shock-paired (A+) or unpaired (B-) cues, for a total of 45 minutes in the session. Each trial was signaled by a common element (X), a 5 s, 1 kHz tone (75 dB) immediately followed by a 15 s discrete auditory (white noise pips) or visible (flashing LED light) CS. Trials were presented in a quasi-random order, so that no cue occurred more than twice in succession. There was a fixed 70 s inter-trial-interval. Assignment of the light or pip as A or B cues was counterbalanced in each experiment, and equally represented in each sex. The Fear Only condition consisted of 15 presentations of the light or PIP (counterbalanced), which co-terminated with shock on the same AX+ trials as the AX+/BX- animals. BX- trials were omitted resulting in an

extended inter-trial-interval so that the conditioning sessions were equal length to the AX+/BX- session. The Control condition received the same auditory and visual stimuli as in AX+/BX- fear discrimination conditioning, but no shocks were presented.

4.2.4 Estrous Phase Testing

Immediately prior to perfusion and tissue collection, all females were tested for estrous phase via vaginal smear with sterile saline. Estrous phase was verified on unstained slides at 10X magnification. Phase verification and procedure were performed as in Cora et al. (2015).

4.2.5 Tissue Collection and Fos Immunohistochemistry Procedures

After conditioning, rats were moved to a quiet room, where they remained undisturbed for 1 hour. One hour after conditioning, all animals were perfused with 0.01M heparinized phosphate buffered saline (PBS) followed by 4% paraformaldehyde. Brains were dissected and post-fixed in 4% paraformaldehyde at 4°C for 24h before being transferred to 30% sucrose. Brains were then sliced into 40µm sections at -20°C and stored in cryoprotectant-filled well plates at 4°C. Immediate early gene product Fos was identified via immunohistochemistry (IHC) as a neural marker of activation.

Fos was visualized as previously, (Rogers-Carter et al., 2018). Free floating sections were blocked with 2% normal donkey serum in PBS-T (0.01%

Triton-X100) and incubated overnight in rabbit anti-c-fos antibody at 1:5000 (Millipore, ABE457). The following morning, sections were washed and incubated in biotinylated donkey anti-rabbit secondary antibody at 1:200 (Jackson ImmunoResearch). Secondary was visualized using the avidin-biotin complex method (ABC Elite Kit, Vector Labs) with chromogen (Vector SG Peroxidase Substrate Kit, Vector Laboratories). At the completion of the reaction, slices were floated onto glass slides, dehydrated, cleared, coverslipped with Permount, and left to dry for 48 hours. Sections were imaged at 10x using a Zeiss Axioimager Z2 light microscope with an AxioCam HRc digital camera. Fos positive cells were quantified within a standardized size area for each region based on atlas images (Paxinos and Watson, 2007). The cell counter plug-in on ImageJ software was used to automate Fos quantification, and parameters were verified by manual cell counts. For each brain region of interest, 2 sections per animal were analyzed.

4.2.6 Brain Regions of Analysis

Brain region	Location from Bregma (mm)	Area analyzed (μm^2)
PL	3.72 to 2.76	570
IL	3.72 to 2.76	460
vlOFC	4.20 to 3.00	570
aIC	4.20 to 3.00	570
mIC	0.48 to -0.24	570
pIC	-1.80 and -2.80	570
BLA	2.04 to -3.36	570
CeA	2.04 to -3.36	190
BNST	0.48 to -0.24	230

Table 4.1 Regions of Fos Analysis

4.2.7 Data Analysis

Freezing was analyzed as percent time freezing during the relative cue and as percentage during the entire 45-minute session. A discrimination index was calculated as freezing to B divided by freezing to A times 100, so that a value of 100 reflected no discrimination between A and B, and values less than 100 indicated reduced fear to B compared to A, as previously (Foilb et al., 2018). Discrimination index was used in correlation analysis with Fos data to explore how well neural activation predicts discrimination.

Group differences in behavioral freezing and Fos were evaluated by analyses of variance (ANOVA) with sex and condition as a between-subjects factors and cue as a within-subjects factor. Main effects and interactions were deemed significant with $p < 0.05$ and between-subjects post hoc comparisons were made with Sidak's correction. Correlations between Fos and discrimination index were analyzed with Pearson's r, as were correlations between Fos in various brain regions. Discrimination index and Fos counts in each brain region were checked for outliers with Grubb's test with alpha set to 0.05. All analyses were made using GraphPad Prism 8.

Due to the lack of sex differences in our analyses of average Fos per region, region-to-region correlations were only performed looking at all animals per condition, both sexes combined. Notably, we did not correct for multiple comparisons in looking at correlations between regions. Since these analyses were largely exploratory, we chose not to correct for multiple comparisons, to allow for the maximum number of caparisons with the potential for more discovery. However making corrections would alter our conclusions and interpretations throughout the chapter, which is important to note as the implications of these findings are discussed. Using Bonferroni correction, a conservative approach, only p-values less than 0.0014 would be considered significant findings in the correlation matrices of regions of interest. This would make only the correlation between PL and IL in the fear condition significant (p = 0.0004). Using a false discovery rate (FDR) correction of 5%, a less conservative

approach, would uphold only the most significant findings of each correlation matrix (Burger, 2018). Using this correction, only the correlations between PL/IL and mIC/BLA would remain significant in the Discrimination condition. In the Fear Only condition, correlations between PL/IL, mIC/vlOFC, and PL/vlOFC remain significant. And finally, in the Control condition, PL/IL and IL/vlOFC correlations remain significant, while the correlation between PL/vlOFC falls just outside of the accepted significant values. We did not correct for multiple comparisons to maximize hypotheses for future directions and combat a somewhat underpowered study for these levels of analysis (elaborated on further in the discussion). Nonetheless, interpretations made using this statistical approach should be made with caution.

4.3 Results

4.3.1 Behavioral Results

4.3.1.1 AX+/BX- Fear Discrimination Conditioning

One male rat from the Discrimination Condition, with significantly greater fear to the B cue compared to the A cue, was identified as an outlier and excluded from all analyses. Sex differences were apparent in AX+/BX- fear discrimination conditioning, as previously (Foilb et al., 2018). Looking at average freezing to each cue during AX+/BX- fear discrimination conditioning (Figure 4.2A), a two-way ANOVA revealed main effects of both cue, $F(1, 13) = 52.64$, $p < 0.0001$, and sex, $F(1, 13) = 12.42$, $p = 0.0037$, but no cue by sex interaction, F

(1, 13) = 0.74, $p = 0.40$. Post hoc analyses showed that males (n = 7) and females (n = 8) displayed significantly different freezing to both cues A, $p = 0.02$, and B, $p = 0.003$. Both males and females significantly discriminated between A and B cues, $p = 0.0015$ and $p < 0.0001$, respectively. To better compare discrimination behavior, a discrimination index was calculated as freezing to B divided by freezing to A multiplied by 100, where a discrimination index of 100 indicates equal freezing to A and B, and lower discrimination indices represent a greater reduction in fear to safe cue B. Females displayed significantly lower discrimination index, representative of greater discrimination, compared to males, t (13) = 2.58, $p = 0.023$ (Figure 4.2B). This trend was evident early in conditioning. A 3-way within subjects ANOVA of sex by cue by trial block (5 blocks of 3 cue trials) identified main effects of cue, F (1,13) = 52.64, $p < 0.0001$, trial block, F (4, 52) = 5.4, $p = 0.001$, sex, F (1,13) = 12.42, $p = 0.0037$, and a significant trial by sex interaction, F (4, 52) = 2.69, $p = 0.04$ (Figure 4.2C). Post hoc comparisons revealed that females significantly discriminated between A and B cues on trial blocks 2 and 3, ps < 0.02, while males did not significantly discriminated between A and B on any trial blocks. Males and females displayed significantly different freezing to B on trial block 2, $p = 0.016$. To investigate a role of estrous phase, we performed a one-way ANOVA, which revealed no main effect of cycle phase on discrimination, F (3, 4) = 1.85, $p = 0.28$. Importantly, since cycle phase was not the main focus of this work, this analysis is underpowered, with only 1 to 3 animals in each of the 4 cycle phases.

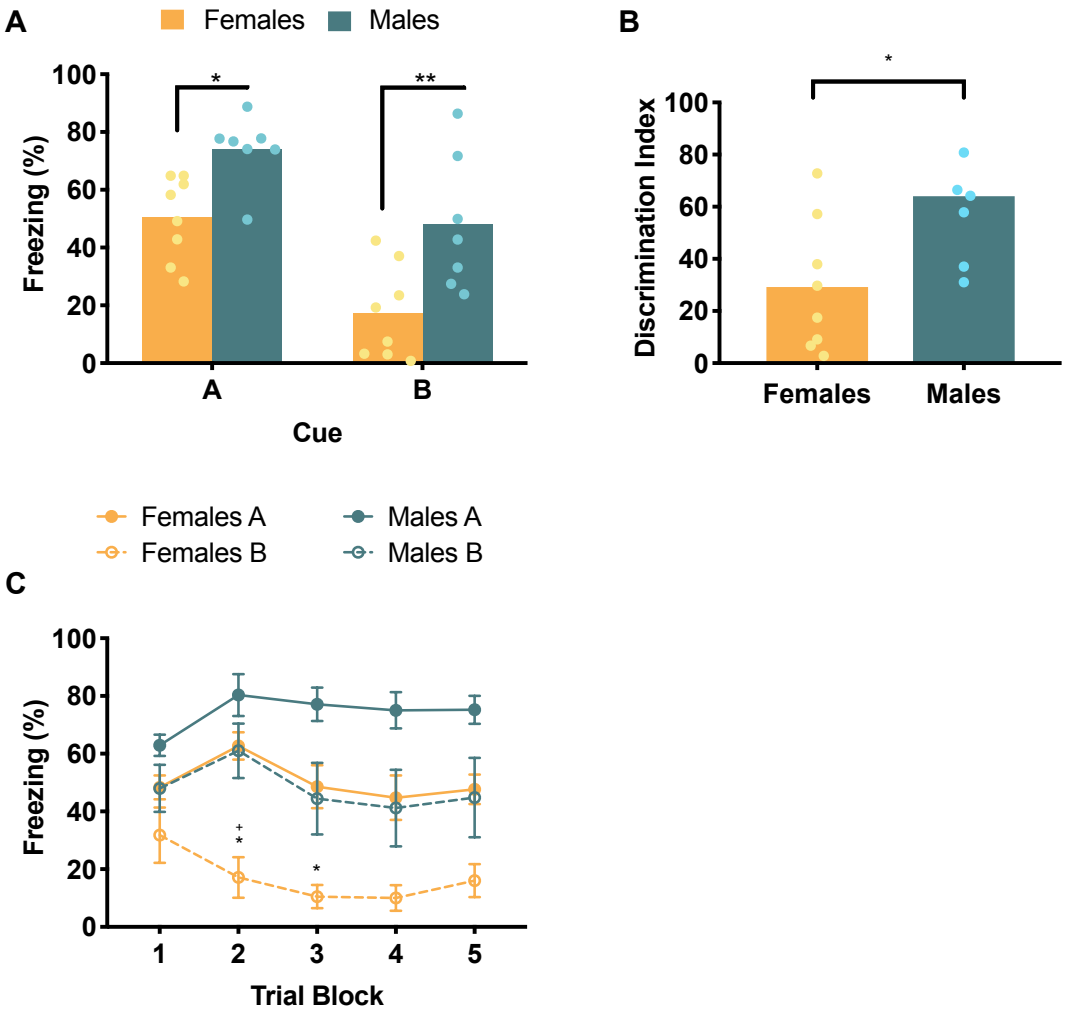

Figure 4.2 Sex differences in fear discrimination behavior. **(A)** Mean (and individual replicates) freezing to A and B cues during AX+/BX- fear discrimination conditioning. Both males and females significantly discriminated between A and B (significance not marked, *p*s < 0.0015). Females froze significantly less than males to both A ($p = 0.02$) and B ($p = 0.003$). **(B)** Mean (and individual replicates) discrimination indices (time freezing to B / time freezing to A x 100) during conditioning. Females had a significantly reduced discrimination index compared to males ($p = 0.023$). **(C)** Freezing averages (±SEM) to A and B across blocks of three trials in conditioning. Females significantly discriminated between A and B in trial blocks 2 and 3 (*p*s < 0.02), while males did not significantly discriminate on any trial blocks. Females also displayed less freezing to B than males on trial block 2 ($p = 0.016$). *significant difference between A and B within sex, $^{+}$significant difference between males and females to B.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$

4.3.1.2 AX+ Fear Conditioning

In the Fear Only condition, a two-tailed t-test comparing males (n = 8) and females (n = 8) average freezing to the danger cue A found trending, but not significant differences between males and females, t (14) = 2.07, $p = 0.06$, which is similar to the effect found on cue A in the Discrimination condition (Figure 4.3A). Despite the variation of fear in females, this did not appear to depend on estrous phase. No females in the fear condition were in estrus. Comparing fear to A in females in proestrous, diestrus, and metestrus cycle phases in a one-way ANOVA, there was no main effect of estrous phase, F (2, 5) = 0.31, $p = 0.75$. Freezing to A cue presentations during conditioning across trial blocks, as in the Discrimination condition, a 2-way ANOVA of sex by trial block found a main effect of trial block, F (4, 56) = 5.47, $p = 0.0009$, but no significant main effect of sex, F (1, 14) = 4.47, $p = 0.053$, and no sex by trial block interaction, F (4, 56) = 0.81, $p = 0.53$ (Figure 4.3B). Post hoc analyses showed significantly different freezing during trial blocks 1 ($p = 0.0004$), 4 ($p = 0.024$), and 5 ($p = 0.025$) compared to trial block 2, where freezing levels peaked.

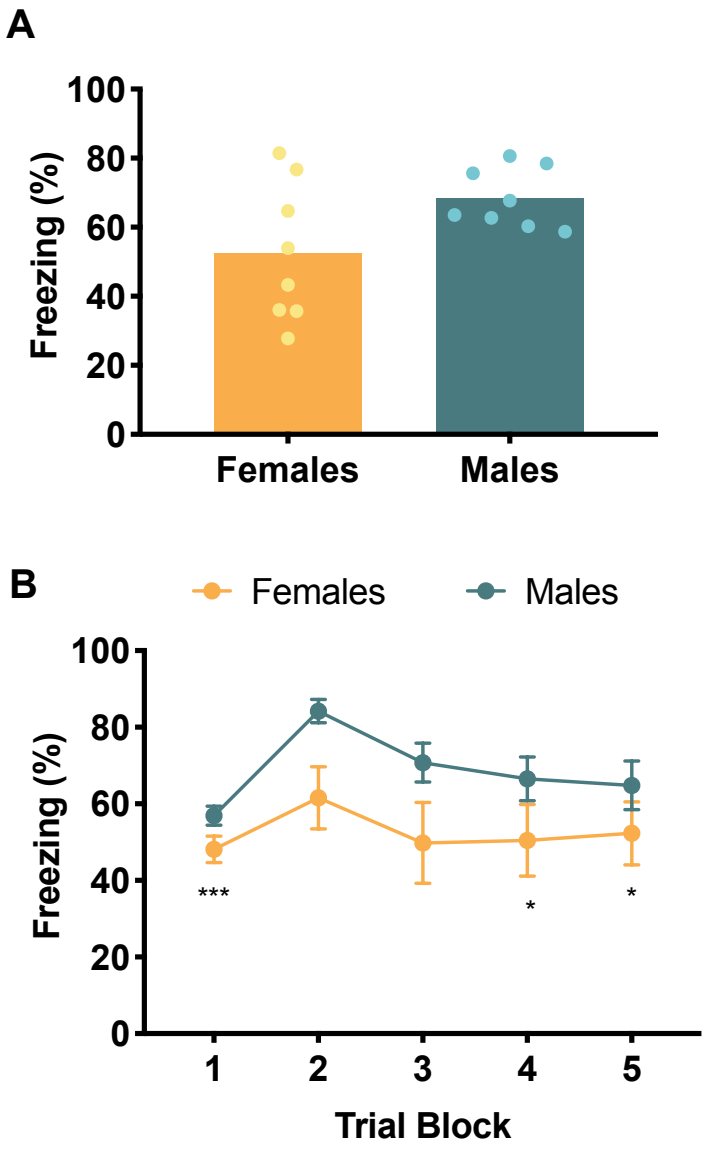

Figure 4.3 Fear conditioning behavior. **(A)** Mean (with individual replicates) fear to danger cue A for animals that underwent AX+ Fear Conditioning. Males and females did not significantly differ in freezing to A ($p = 0.06$). **(B)** Mean (±SEM) to A across blocks of three trials in conditioning. Freezing was significantly different in trial blocks 1, 4 and 5 compared to trial 2 (ps <0.025). * $p < 0.05$, *** $p < 0.001$

4.3.1.3 All Behavioral Conditions

Analysis of freezing time during the 45 minute conditioning session for all behavioral conditions, a 2-way ANOVA found main effects of condition, $F(2, 41) = 131.1$, $p < 0.0001$, and sex, $F(1, 41) = 8.36$, $p = 0.006$, but no condition by sex interaction, $F(2, 41) = 1.42$, $p = 0.25$ (Figure 4.4). As expected, post hoc analyses revealed significantly increased freezing in Fear Only and Discrimination conditions compared to Control animals (ps < 0.0001). This result indicates that our Control animals did not express freezing to the cues or conditioning context. Males and females significantly differed in overall time freezing in the Discrimination condition, $p = 0.015$, but not in Fear Only or Control conditions, ps > 0.05. This further supports the importance of exploring sex differences in discrimination learning.

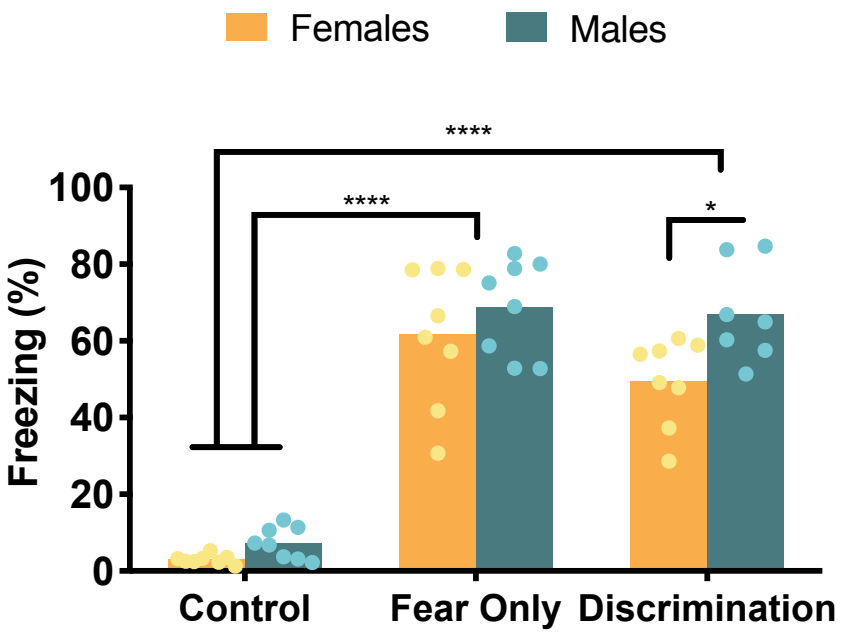

Figure 4.4 Freezing behavior in all conditions. Mean (and individual replicates) percent time freezing during the entire 45 minute conditioning session for all behavior conditions. Control animals froze significantly less than both Fear Only and Discrimination animals (*p*s < 0.0001). In the Discrimination condition, females displayed less total freezing time compared to males ($p = 0.015$).

4.3.2 Fos Results

Brain Region	Condition	Sex	Interaction
PL	**F (2, 40) = 8.66*** **	F (1, 40) = 1.14	F (2, 40) = 0.23
IL	**F (2, 40) = 4.85***	F (1, 40) = 0.08	F (2, 40) = 0.41
vlOFC	F (2, 39) = 0.97	F (1, 39) = 1.24	F (2, 39) = 0.23
aIC	F (2, 41) = 1.68	F (1, 41) = 0.16	F (2, 41) = 1.34
mIC	**F (2, 40) = 3.54***	F (1, 40) = 0.0009	F (2, 40) = 0.73
pIC	F (2, 41) = 1.70	F (1, 41) = 0.65	F (2, 41) = 1.68
BLA	**F (2, 41) = 9.02*** **	F (1, 41) = 1.46	F (2, 41) = 0.79
CeA	**F (2, 41) = 19.64**** **	F (1, 41) = 0.017	F (2, 41) = 2.36
BNST	**F (2, 39) = 13.56**** **	**F (1, 39) = 6.59***	F (2, 39) = 0.53

Table 4.2 Brain Region Fos by Condition and Sex. Results of a 2 (sex) by 3 (condition) ANOVA of average Fos counts for each brain region. Regions in bold represent significant main effects. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$

Brain Region	All Animals	Female	Males
PL	r = -0.47, $p = 0.077^{\#}$	r = -0.21, $p = 0.617$	r = -0.65, $p = 0.116$
IL	r = -0.37, $p = 0.174$	r = -0.40, $p = 0.329$	r = -0.40, $p = 0.374$
vlOFC	r = 0.069, $p = 0.808$	r = 0.15, $p = 0.727$	r = 0.16, $p = 0.739$
aIC	r = 0.27, $p = 0.340$	r = -0.13, $p = 0.758$	r = 0.019, $p = 0.968$
mIC	r = -0.14, $p = 0.631$	r = -0.39, $p = 0.342$	r = 0.36, $p = 0.433$
pIC	r = -0.29, $p = 0.290$	r = -0.25, $p = 0.552$	r = 0.19, $p = 0.689$
BLA	r = -0.06, $p = 0.837$	r = 0.24, $p = 0.567$	r = 0.30, $p = 0.512$
CeA	**r = -0.59, $p = 0.019$***	r = -0.63, $p = 0.091^{\#}$	r = -0.50, $p = 0.256$
BNST	r = 0.13, $p = 0.659$	r = -0.65, $p = 0.084^{\#}$	r = 0.11, $p = 0.843$

Table 4.3 Brain Region Fos and Discrimination Correlations. Pearson's r correlation (and corresponding p-value) between brain region and discrimination index across all animals in the Discrimination condition, and by sex. Regions in bold represent significant correlations. # $p < 0.1$ (trending), * $p < 0.05$ (significant).

4.3.2.1 Prefrontal Cortex

4.3.2.1.1 PL

One female in the Control condition was not analyzed for PL Fos due to damaged brain slices. Therefore, analysis of PL Fos included: Control Females (n = 7), Control Males (n = 8), Fear Only Females (n = 8), Fear Only Males (n = 8), Discrimination Females (n = 8), Discrimination Males (n = 7). Analysis of PL Fos revealed a main effect of condition, but no main effect of sex or sex by condition interaction (see Table 4.2). Post hoc comparisons found that PL Fos was significantly increased in Fear Only (p = 0.023) and Discrimination (p =

0.0002) conditions compared to Control animals, while Fos in Fear Only and Discrimination conditions did not significantly differ, but was trending toward significance ($p = 0.07$), with increased Fos in the Discrimination condition compared to Controls (Figure 4.5B). Examining the Discrimination condition more closely, we aimed to uncover if Fos in any brain region of interest directly correlated with animals' ability to discrimination between A and B cues, as measured by the discrimination index. Pearson r correlation was calculated between discrimination index and PL Fos (Table 4.3). Across all Discrimination animals, PL Fos and discrimination index trended toward a significant correlation ($p = 0.077$), where there were higher levels of Fos in PL in animals with lower discrimination indices – demonstrating greater discrimination (Figure 4.5C). Representative PL Fos from each condition is displayed in Figure 4.6.

4.3.2.1.2 IL

As in PL, brain sections from one female from the Control condition was unable to be analyzed for IL Fos. Analysis of average IL Fos revealed a main effect of condition, but no main effect of sex or sex by condition interaction (Table 4.2), with reduced Fos in Control animals compared to both Fear Only ($p = 0.04$) and Discrimination ($p = 0.004$) conditions (Figure 4.5E). There was no significant difference between Fear Only and Discrimination conditions ($p = 0.32$). Pearson r correlation did not find a significant correlation between IL Fos and discrimination

index (Figure 4.5F; Table 4.3). Representative images of IL Fos are displayed in Figure 4.7.

4.3.2.1.3 OFC

vlOFC Fos counts for one male in the Control condition was identified as an outlier by Grubbs' test, alpha set to 0.05, and was therefore excluded from analyses. The resulting groups were as follows: Control Females (n = 8), Control Males (n = 7), Fear Only Females (n = 8), Fear Only Males (n = 8), Discrimination Females (n = 8), Discrimination Males (n = 7). Unlike the subregions of vmPFC, no significant main effects were found in average vlOFC Fos (Table 4.2; Figure 4.5H). There was also no significant correlation between vlOFC Fos and discrimination index (Table 4.3; Figure 4.5I).

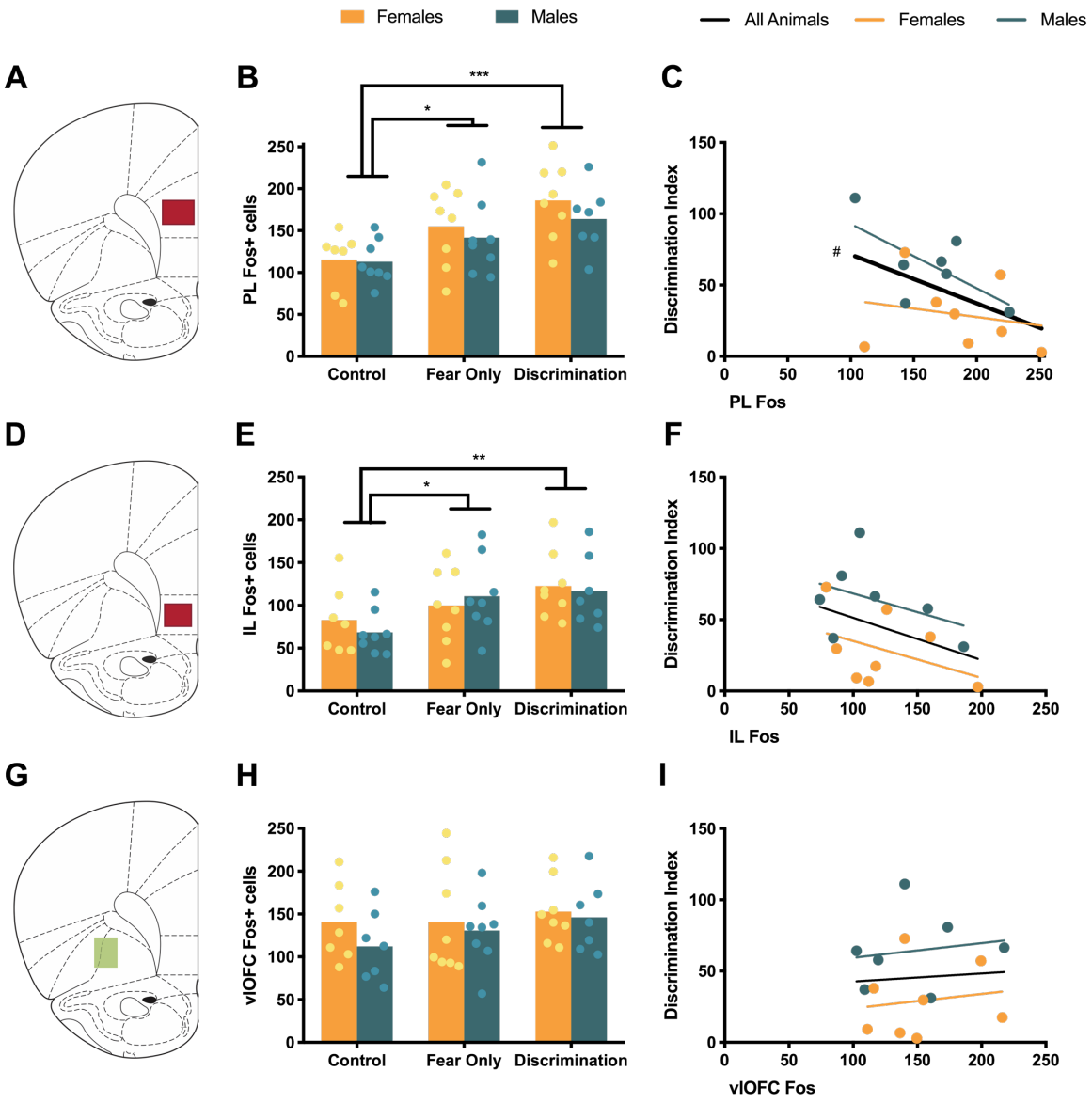

Figure 4.5 Fos in Prefrontal Cortex. **(A)** Atlas image at +3.20 mm from Bregma with representative PL area of analysis in red. **(B)** Mean (with individual replicates) Fos positive cells in the PL. **(C)** Correlation between PL Fos and discrimination index, # indicates trending effect, $p = 0.077$. **(D)** Atlas image at +3.20 mm from Bregma with representative IL area of analysis in red. **(E)** Mean (with individual replicates) Fos positive cells in the IL. **(F)** Correlation between IL Fos and discrimination index. **(G)** Atlas image at +3.20 mm from Bregma with representative vlOFC area of analysis in green. **(H)** Mean (with individual replicates) Fos positive cells in the vlOFC. **(I)** Correlation between vlOFC Fos and discrimination index. $^{\#}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $p < 0.001$

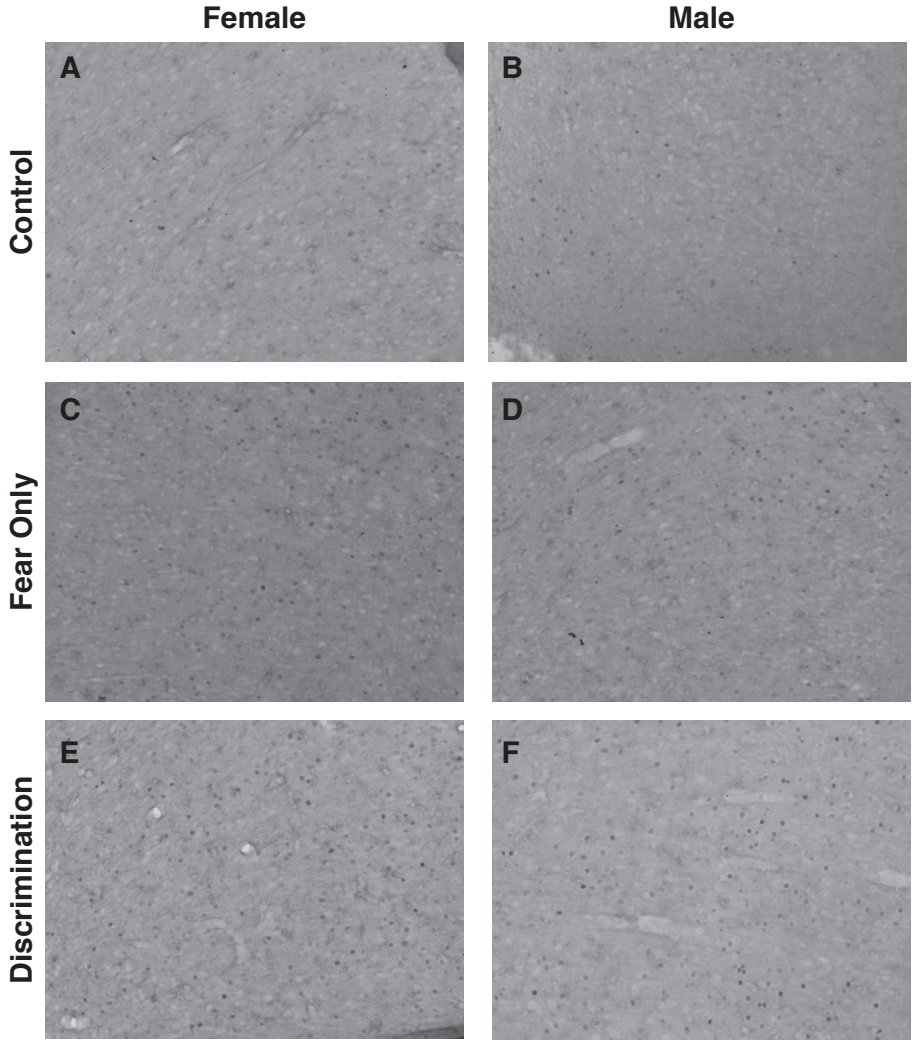

Figure 4.6 Representative PL Fos. (A) Control Females, **(B)** Control Males, **(C)** Fear Only Females, **(D)** Fear Only Males, **(E)** Discrimination Females, **(F)** Discrimination Males.

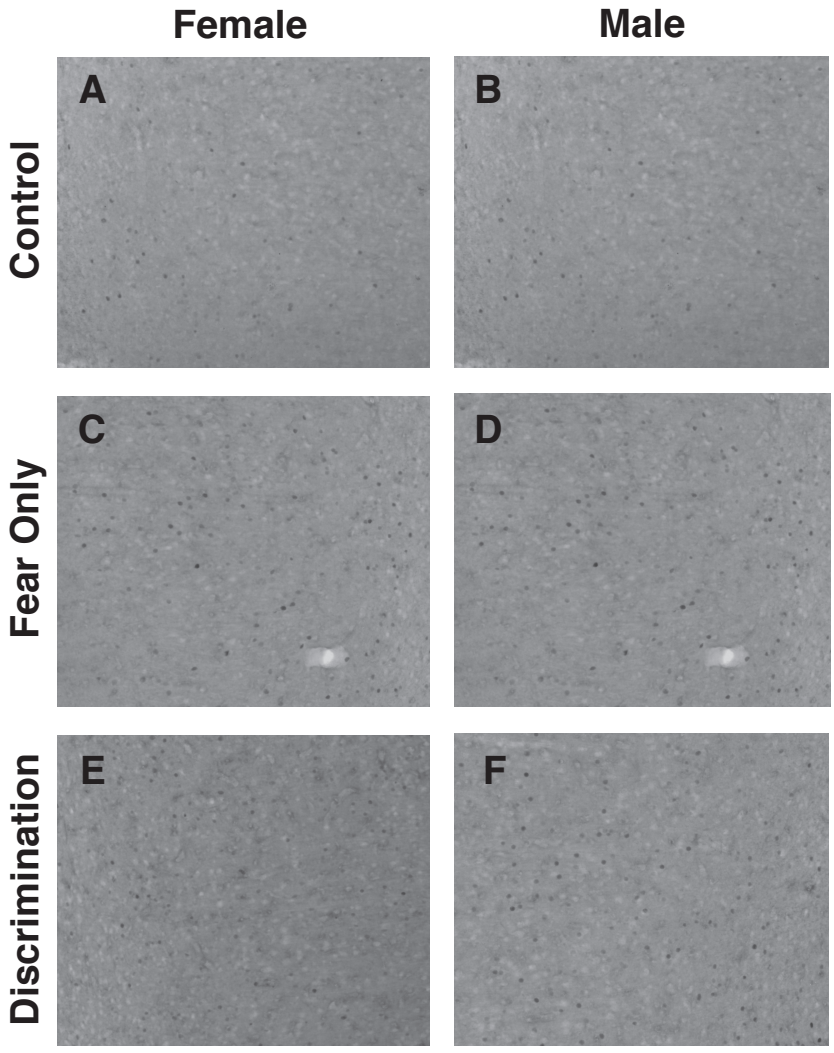

Figure 4.7 Representative IL Fos. (A) Control Females, **(B)** Control Males, **(C)** Fear Only Females, **(D)** Fear Only Males, **(E)** Discrimination Females, **(F)** Discrimination Males.

4.3.2.2 Insular Cortex

4.3.2.2.1 aIC

Analysis of average Fos in aIC revealed no main effects of condition or sex, and no sex by condition interaction (Table 4.2, Figure 4.8B). There was also no significant correlation between aIC Fos and discrimination index (Table 4.3, Figure 4.8C).

4.3.2.2.2 mIC

Fos measurements from mIC were not obtained for one female in the Control condition due to damaged brain slices. Therefore, analysis of mIC Fos included: Control Females (n = 7), Control Males (n = 8), Fear Only Females (n = 8), Fear Only Males (n = 8), Discrimination Females (n = 8), Discrimination Males (n = 7). A main effect of condition, but no main effect of sex or sex by condition interaction, was found in average Fos measures in mIC (Table 4.2). Post hoc analysis revealed significantly different mIC Fos in Control and Discrimination groups ($p = 0.011$), but not Fear Only compared to Discrimination animals ($p = 0.2$) or Controls ($p = 0.16$; Figure 4.8E). There was no significant correlation between mIC Fos and discrimination Index (Table 4.3; Figure 4.8F).

4.3.2.2.3 pIC

As in aIC, no main effects of condition or sex, or a sex by condition interaction were found in analysis of average pIC Fos (Table 4.2; Figure 4.8H).

There were also no significant correlations between pIC Fos and discrimination index (Table 4.3; Figure 4.8I).

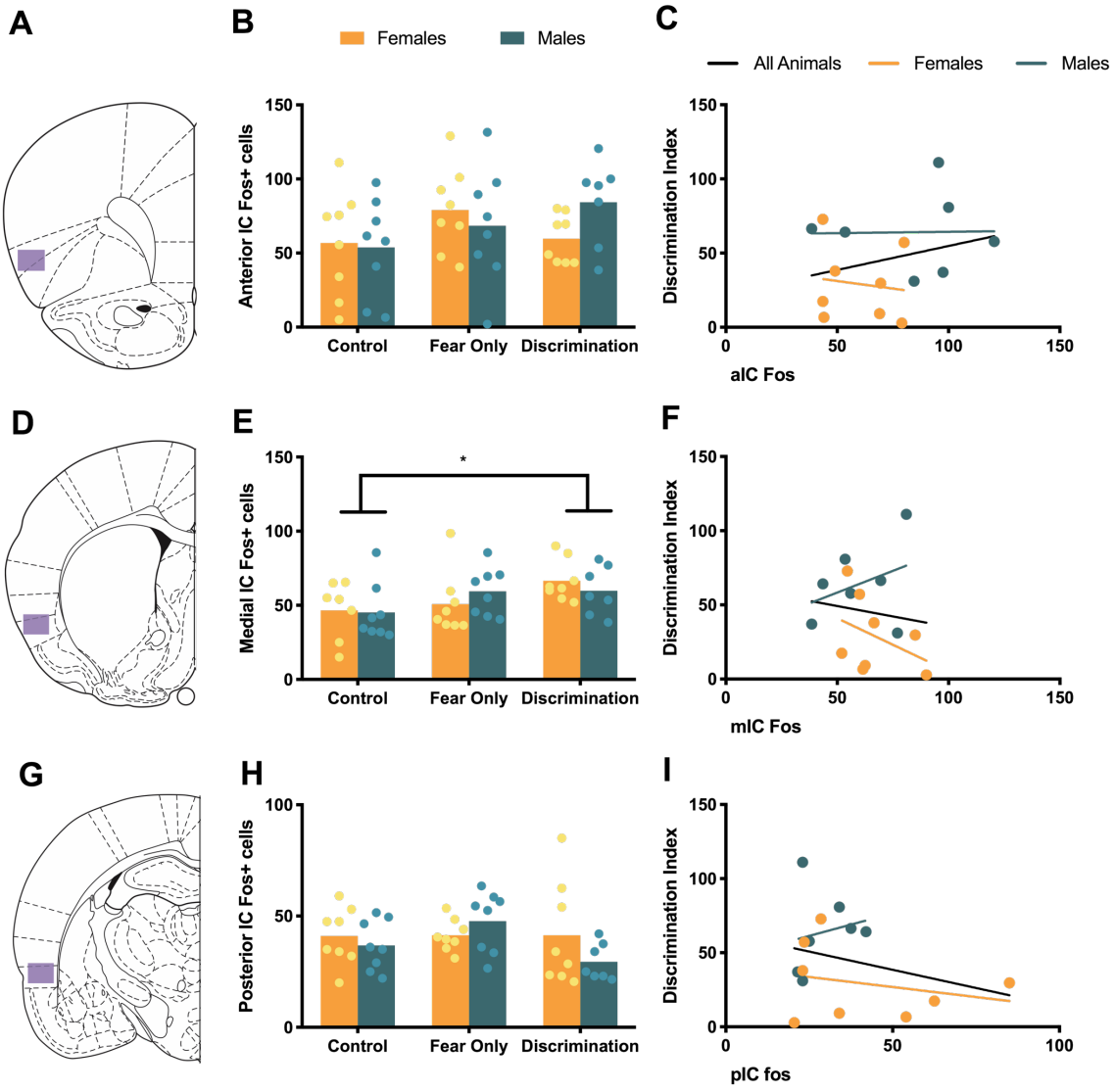

Figure 4.8 Fos in Insular Cortex. (A) Atlas image at +3.20 mm from Bregma with representative aIC area of analysis in purple. **(B)** Mean (with individual replicates) Fos positive cells in aIC. **(C)** Correlation between aIC Fos and discrimination index. **(D)** Atlas image at +0.20 mm from Bregma with representative mIC area of analysis in purple. **(E)** Mean (with individual replicates) Fos positive cells in mIC. **(F)** Correlation between mIC Fos and discrimination index. **(G)** Atlas image at -2.56 mm from Bregma with representative pIC area of analysis in purple. **(H)** Mean (with individual replicates) Fos positive cells in the pIC. **(I)** Correlation between pIC Fos and discrimination index.
* $p < 0.05$

4.3.2.3 Extended Amygdala

4.3.2.3.1 BLA

Analysis of BLA Fos revealed a main effect of condition, but no main effect of sex and no sex by condition interaction (Table 4.2). Post hoc analyses showed significantly increased Fos positive cells in BLA in Fear Only ($p = 0.001$) and Discrimination ($p = 0.0004$) conditions compared to Control animals (Figure 4.9B). Pearson r correlation analysis of BLA Fos and discrimination index found no significant correlations (Table 4.3; Figure 4.9C). Representative images of BLA Fos are displayed in Figure 4.10.

4.3.2.3.2 CeA

Analysis of CeA was biased towards lateral CeA (CEl) and capsular CeA (CEc), although medial CeA (CEm) was not specifically excluded from analysis. As in BLA, CeA Fos analysis revealed a main effect of condition, with significantly increased Fos in shocked conditions – Fear Only and Discrimination (ps < 0.0001) – compared to Control animals (Figure 4.9E), but no significant difference in Fos between Fear Only and Discrimination conditions ($p = 0.43$). No main of effect of sex or interaction of sex and condition was found in CeA Fos (Table 4.2). A significant correlation between CeA Fos and discrimination index was apparent, Pearson r = -0.59, $p = 0.02$, such that higher CeA Fos indicates a lower discrimination index, or greater discrimination between A and B cues (Figure 4.9F). CeA representative Fos are displayed in Figure 4.11.

4.3.2.3.3 BNST

Damaged tissue slices for BNST prevented Fos collection from one Control female and one Discrimination Male. Fos analysis of BNST included: Control Females (n = 7), Control Males (n = 8), Fear Only Females (n = 8), Fear Only Males (n = 8), Discrimination Females (n = 8), Discrimination Males (n = 6). Anterior medial BNST Fos was the only brain region of our selected regions of interest where we found a significant main effect of sex. There was also a significant main effect of condition in BNST Fos, but no sex by condition interaction (Table 4.2). Post hoc analyses revealed significantly less Fos positive cells in females of the Discrimination condition compared to males in the Discrimination condition ($p = 0.04$), but no sex differences were found in Fear Only ($p = 0.14$) or Control ($p = 0.45$) conditions. Between conditions, there was significantly increased BNST Fos in the Fear Only condition compared to Controls ($p = 0.027$) and in Discrimination compared to both the Control ($p < 0.0001$) and Fear Only ($p = 0.004$) conditions (Figure 4.9H). Discrimination index did not correlate with BNST Fos (Table 4.3). However the correlation between BNST Fos and discrimination in female discrimination animals trended towards significance ($p = 0.08$), such that more BNST Fos indicated a lower discrimination index, or greater discrimination between cues (Figure 4.9I). Representative BNST Fos sections are displayed in Figure 4.12.

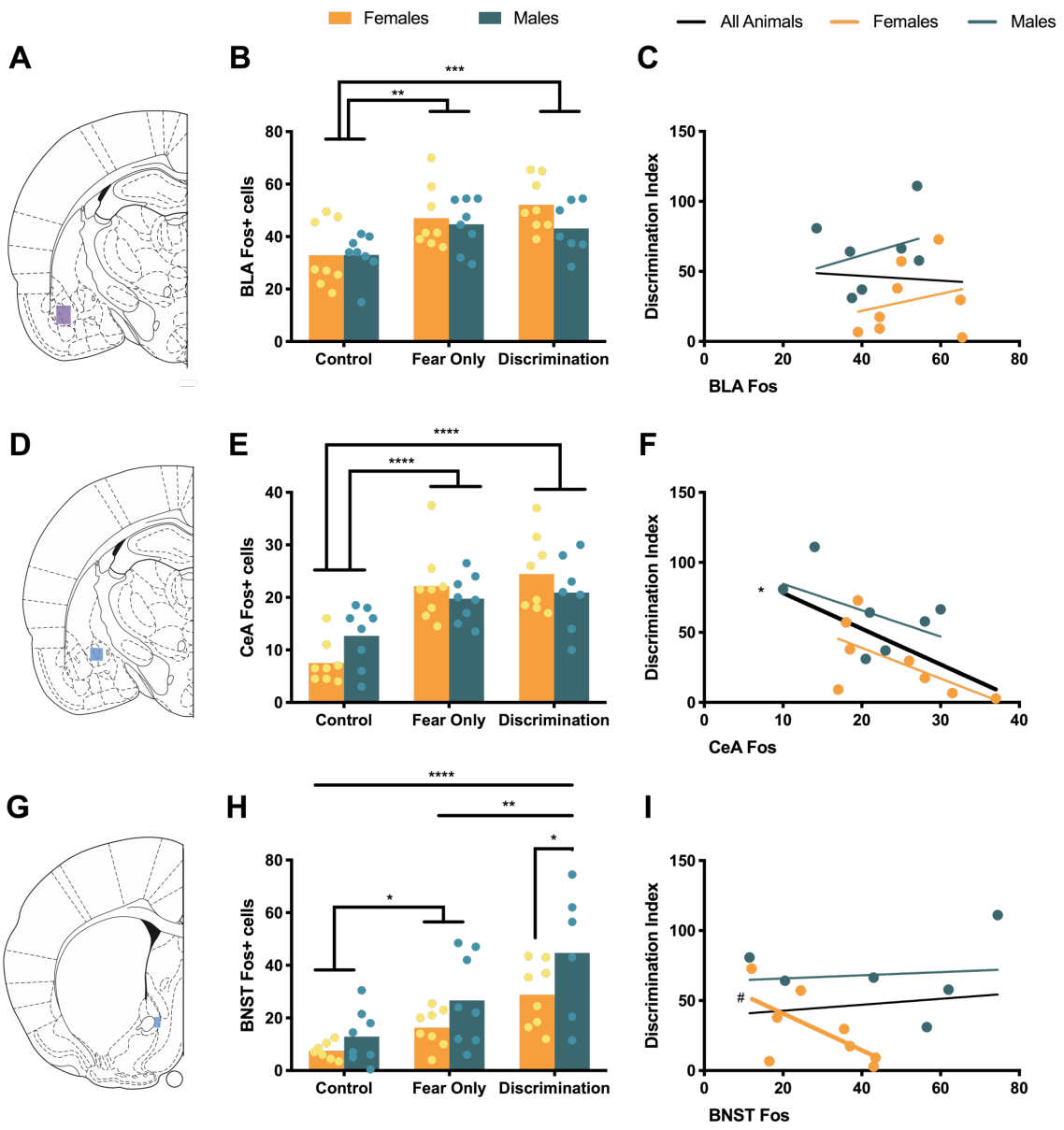

Figure 4.9 Fos in the Extended Amygdala. (A) Atlas image at -2.56 mm from Bregma with representative BLA area of analysis in purple. **(B)** Mean (with individual replicates) Fos positive cells in BLA. **(C)** Correlation between BLA Fos and discrimination index. **(D)** Atlas image at -2.56 mm from Bregma with representative CeA area of analysis in blue. **(E)** Mean (with individual replicates) Fos positive cells in CeA. **(F)** Correlation between CeA Fos and discrimination index. **(G)** Atlas image at +0.20 mm from Bregma with representative anterior medial BNST area of analysis in blue. **(H)** Mean (with individual replicates) Fos positive cells in the BNST. **(I)** Correlation between BNST Fos and discrimination index. $^{\#}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, $^{****}p < 0.0001$

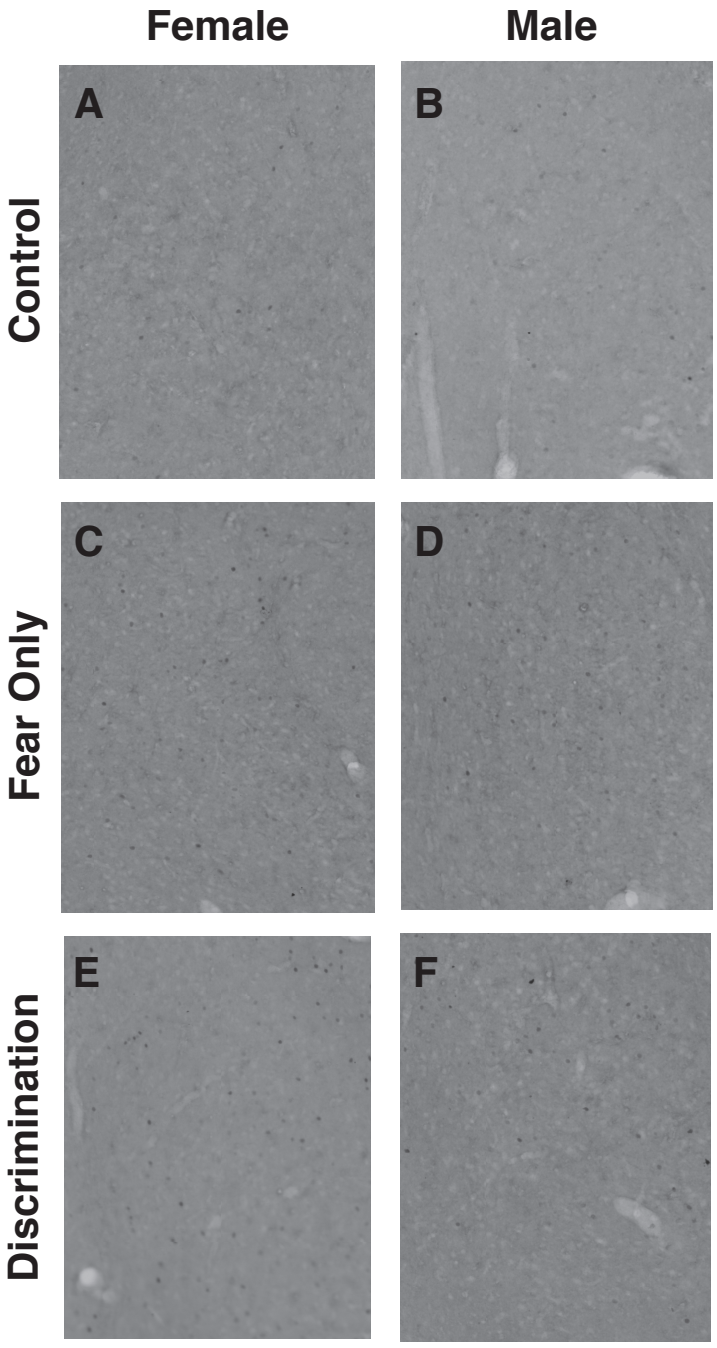

Figure 4.10 Representative BLA Fos. (A) Control Females, **(B)** Control Males, **(C)** Fear Only Females, **(D)** Fear Only Males, **(E)** Discrimination Females, **(F)** Discrimination Males.

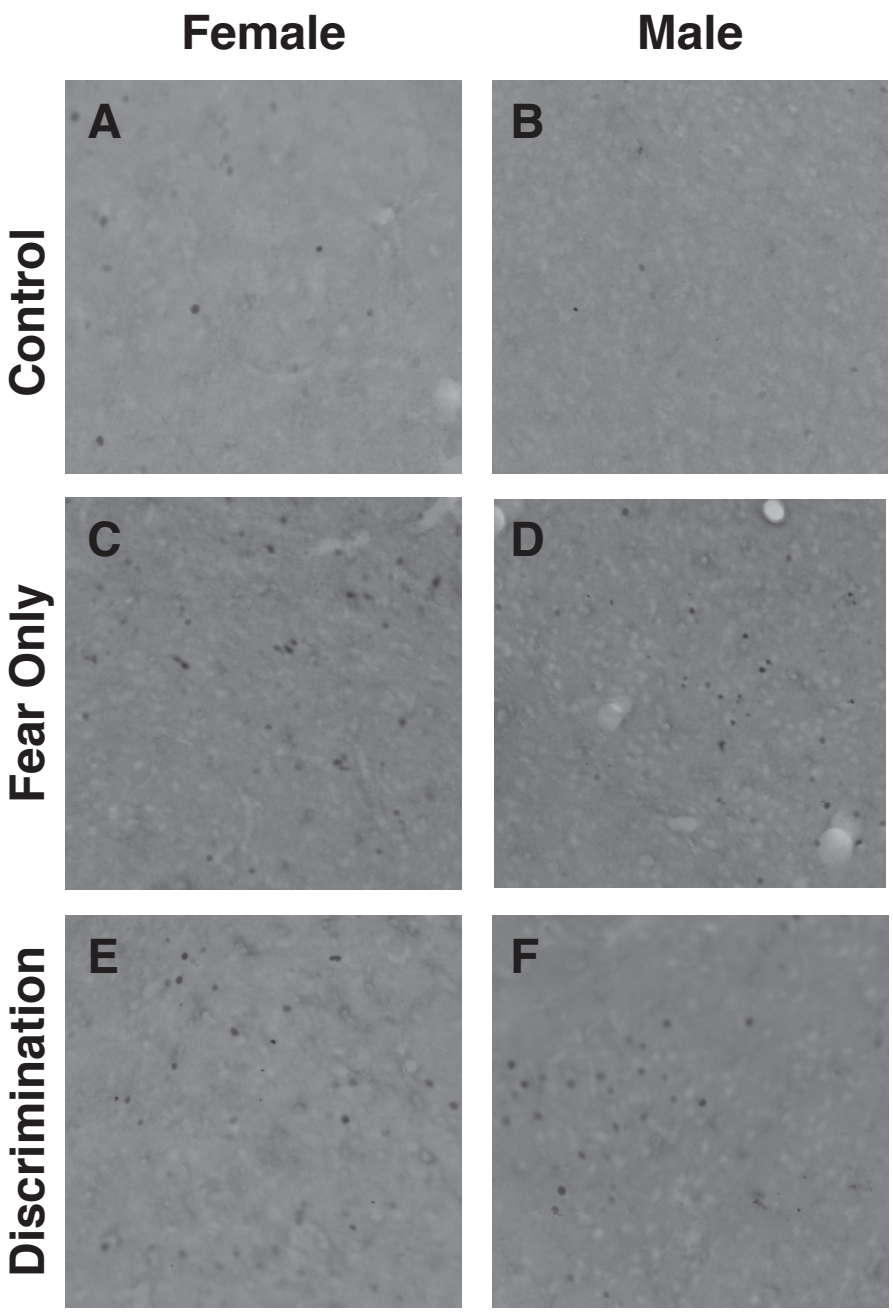

Figure 4.11 Representative CeA Fos. (A) Control Females, **(B)** Control Males, **(C)** Fear Only Females, **(D)** Fear Only Males, **(E)** Discrimination Females, **(F)** Discrimination Males.

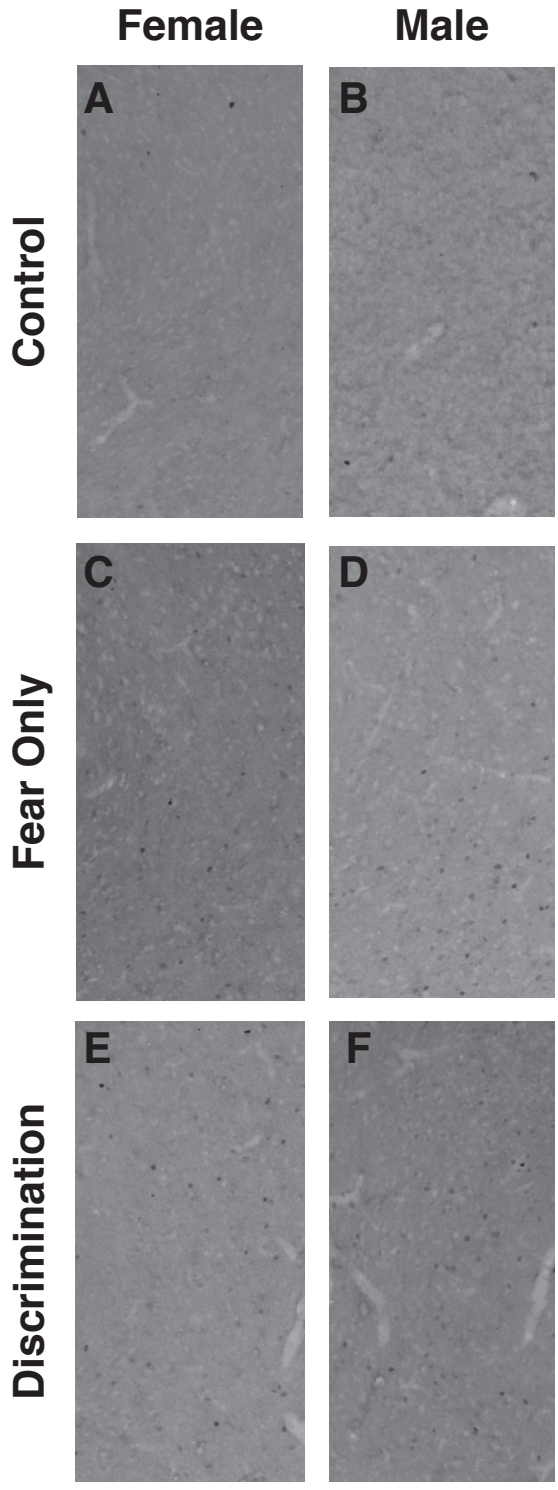

Figure 4.12 Representative BNST Fos. (A) Control Females, **(B)** Control Males, **(C)** Fear Only Females, **(D)** Fear Only Males, **(E)** Discrimination Females, **(F)** Discrimination Males.

4.3.2.4 Region Connectivity

Correlations between brain regions were used as a measure of functional connectivity. Due to a lack of sex differences in all regions aside from BNST, functional connectivity was analyzed as a combination of all males and females in each condition. Correlation matrices for the Control condition (Table 4.4), Fear Only condition (Table 4.5) and Discrimination condition (Table 4.6) are represented in Figure 4.13. In all conditions, Fos in PL and IL significantly correlated (ps < 0.02). In the Control condition, vlOFC also significantly correlated with both PL and IL (ps < 0.02). The correlation between vlOFC and PL was also significant in the Fear Only condition (p = 0.019). IL and mIC significantly correlated in the Discrimination condition (p = 0.04). Fos in BLA was a particularly interesting because it correlated with different subregions of IC in each condition: with aIC in Controls (p = 0.016), with pIC in Fear Only (p = 0.032) and with mIC in Discrimination (p = 0.019).

Mean correlation coefficients were calculated for each condition from the correlation matrices to compare functional connectivity of the brain regions analyzed (Wheeler et al., 2013). A one-way ANOVA comparing the mean correlation coefficients in Control, Fear Only and Discrimination conditions found no significant differences in functional connectivity across the three conditions, $F(2, 105) = 1.115$, $p = 0.33$.

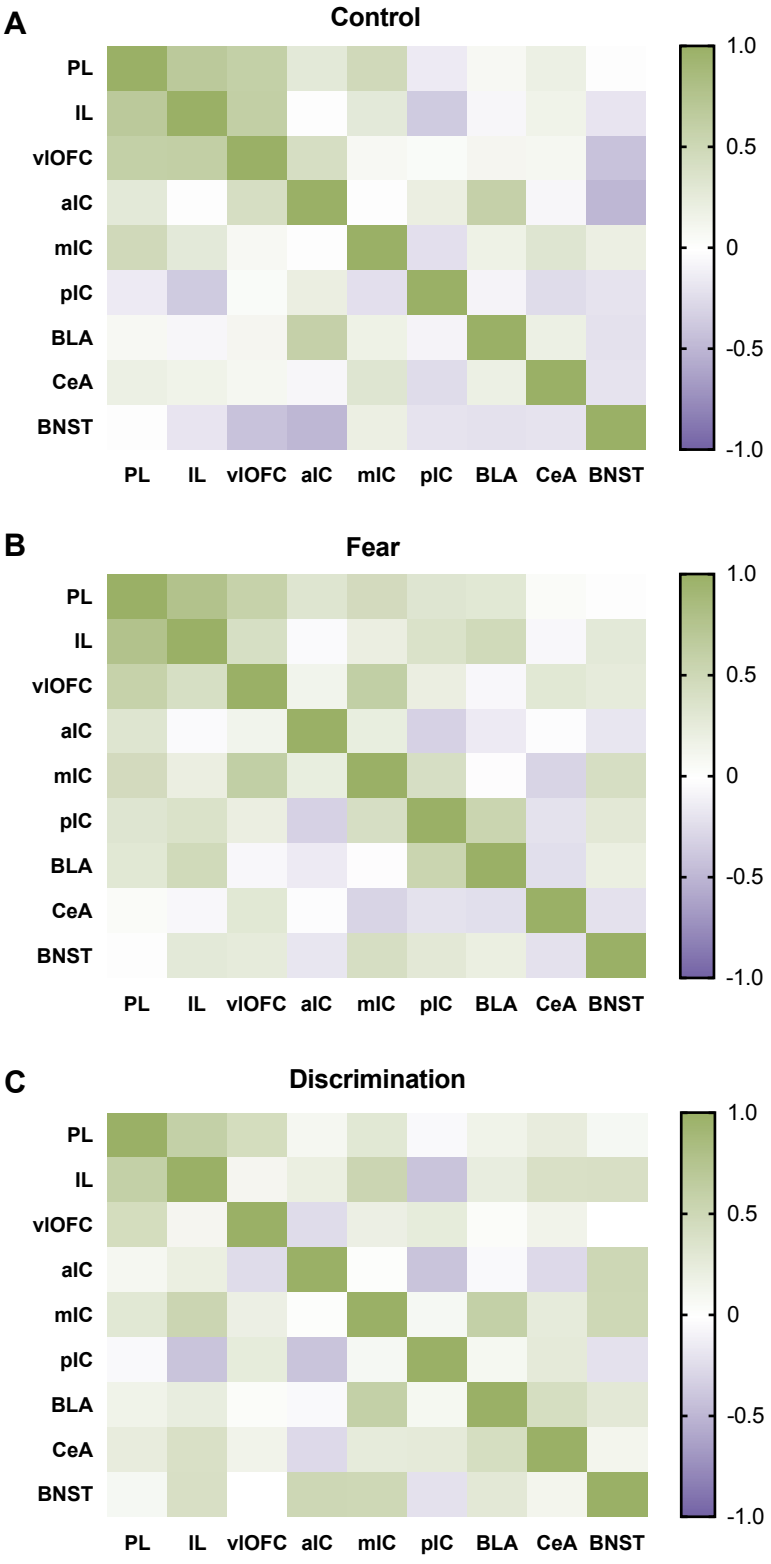

Figure 4.13 Fos correlations between regions. Visualization of correlation matrices displayed in Tables 4.3 **(A)** Control, 4.4 **(B)** Fear Only, and 4.5 **(C)** Discrimination. Green squares indicate a positive r value, while purple values signify a negative relationship.

	PL	IL	vlOFC	aIC	mIC	pIC	BLA	CeA
PL								
IL	**0.69****							
vlOFC	**0.60***	**0.62***						
aIC	0.27	0.01	0.42					
mIC	0.48	0.27	0.08	0.01				
pIC	-0.14	-0.36	0.05	0.21	-0.22			
BLA	0.08	-0.07	0.11	**0.59***	0.17	-0.08		
CeA	0.19	0.16	0.10	-0.06	0.32	-0.25	0.19	
BNST	0.01	-0.18	-0.42	-0.48	0.20	-0.19	-0.21	-0.19

Table 4.4 Region Correlations in the Control Condition. Correlations (r values) of Fos across brain regions for animals in the control conditions. *$p < 0.05$, **$p < 0.01$

	PL	IL	vlOFC	aIC	mIC	pIC	BLA	CeA
PL								
IL	**0.78*** **							
vlOFC	**0.58***	0.41						
aIC	0.33	-0.04	0.14					
mIC	0.46	0.21	**0.62***	0.22				
pIC	0.32	0.38	0.21	-0.32	0.42			
BLA	0.30	0.48	-0.06	-0.15	-0.02	**0.54***		
CeA	0.05	-0.06	0.30	-0.02	-0.30	-0.21	-0.22	
BNST	0.004	0.27	0.25	-0.18	0.42	0.29	0.20	-0.21

Table 4.5 Region Correlations in the Fear Only Condition. Correlations (r values) of Fos across brain regions for animals in the control conditions. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

	PL	IL	vlOFC	aIC	mIC	pIC	BLA	CeA
PL								
IL	**0.60***							
vlOFC	0.44	0.11						
aIC	0.10	0.20	-0.25					
mIC	0.30	**0.53***	0.20	0.03				
pIC	-0.05	-0.41	0.25	-0.41	0.08			
BLA	0.16	0.23	0.04	-0.05	**0.59***	0.10		
CeA	0.24	0.39	0.15	-0.26	0.26	0.26	0.43	
BNST	0.09	0.40	-0.002	0.52	0.51	-0.21	0.29	0.13

Table 4.6 Region Correlations in the Discrimination Condition. Correlations (r values) of Fos across brain regions for animals in the control conditions. *$p < 0.05$

4.4 Discussion

The first goal of this work was to identify brain regions differentially activated by safety learning compared to fear learning and controls, and to explore sex differences in activation in these regions. While sex differences were only found in the BNST, main effects of behavioral condition were found in the number of Fos positive cells in many brain regions of interest. In mIC, there was significantly more Fos in the Discrimination condition than in Control animals and in BNST each of the three behavioral conditions had significantly different numbers of Fos positive cells compared to the other conditions. As expected, many brain regions showed increased neuronal activation to shock conditions (Discrimination and Fear Only) compared to the Control condition, including PL, IL, BLA, and CeA. While the overlap in activation seen in Fear Only and Discrimination may be due to the large overlap in the paradigms, and the presence of fear learning in Discrimination, it is possible that these regions are being activated differently in Discrimination animals compared to Fear Only. For example, different populations of neurons could be activated in the presence of a safety cue. The lack of temporal specificity with Fos also prevents us from determining if any neurons were specifically activated in response to the safety cue. To further determine the relationship between Fos and discrimination, we explored correlations between region activation and discrimination index, and surprisingly found that only Fos in the CeA predicts discrimination abilities. Finally we tested if the functional connectivity of these regions differed based on

experimental treatment. While there was not a significant difference in overall functional connectivity of the brain regions we analyzed, there were clear differences between behavioral conditions, which is informative in understanding the mechanisms that allow the appropriate modulation fear to safety cues.

Here I will review the main findings of this research and discuss implications of these results on the future of safety learning research. In doing so, I outline promising avenues for future investigations on the neural mechanisms that allow for the discrimination between safety and danger. Throughout this discussion I reference limitations of using Fos as a neural marker of activation. Since Fos doesn't identify 1) cell types: different numbers of excitatory and inhibitory neurons may be activated in different behavioral conditions, 2) neural projections: there is ample evidence that activation of specific outputs support specific conditions (Baratta et al., 2009; Do-Monte et al., 2015), 3) or temporal specificity: levels of fear clearly vary throughout conditioning in discrimination learning and Fos is unable to identify which neurons are responding to presentations of the danger cue vs. the safety cue. Nonetheless, the findings of this work provide significant data to draw conclusions on the neural nodes underlying safety learning and begin to indicate the neural pathways that allow for the modulation of fear by a safety signal.

4.4.1 Prefrontal Cortex

Increased Fos was seen in PL and IL regions of the vmPFC in Fear Only and Discrimination conditions compared to Controls. This finding was unsurprising due to the role of these structures in fear (Sotres-Bayon and Quirk, 2010; Sierra-Mercado et al., 2011). We predicted that vmPFC likely plays a role in the processing of danger information in a discrimination paradigm due to prior discrimination research that found reduction of fear to a danger cue after inhibition of either PL or IL and increased excitability of vmPFC in response to presentation of danger cue (Likhtik et al., 2014; Sangha et al., 2014). While Likhtik et al. (2014), found increased synchrony in vmPFC and BLA firing with discrimination learning, we did not find evidence for connectivity between vmPFC and BLA Fos. This is likely due the temporal specificity of their finding, which cannot be observed with Fos. The correlation between PL and discrimination index was trending toward significance, and increased sample size may have also revealed that PL Fos is able to predict discrimination abilities. In the human literature, fMRI results show stronger responses of the vmPFC during acquisition of the safety cue compared to the danger cue, which does not seem to align with the rodent work on vmPFC (Schiller et al., 2008).

There was a significant correlation between PL and IL Fos in all conditions, which is likely a result of their dense interconnectivity (Hoover and Vertes, 2007). Similarly, PL and IL both correlated with vlOFC in the control condition and IL correlates with mIC in the Fear Only condition. These

correlations all fit with the known structural connectivity of vlOFC and IC with medial prefrontal cortex (Hurley et al., 1991; Shi and Cassell, 1998a, b; Hoover and Vertes, 2007).

Our findings support a role of vmPFC in danger processing of fear discrimination. In humans, there is evidence that the impairments in safety signal processing in individuals with PTSD could involve problems with top-down emotional control by the vmPFC (Rauch et al., 2006). Findings of structural differences in the vmPFC are also reported in PTSD populations (Corbo et al., 2005; Etkin and Wager, 2007; Hughes and Shin, 2011). Previous findings and the work presented here indicate that vmPFC is likely engaged with danger learning and could play a role in over-active fear responding, rather than a deficit in utilization of safety information.

On the side of safety processing, our lab and others previously found evidence that vlOFC is necessary for a reduction of fear in the presence of safety signals (Ray et al., 2018; Sarlitto et al., 2018). However there was no effect in vlOFC Fos across conditions or sex, and Fos in vlOFC did not correlate with discrimination index or any other brain regions in the discrimination condition. The lack of findings in OFC are surprising due to its connectivity to other critical fear and discrimination structures. OFC and amygdala share extensive connective similarities, such as extensive limbic connectivity and projections to midbrain regions responsible for behavioral output (Swanson and Petrovich,

1998; LeDoux, 2000; Price, 2007). OFC also projects to IC and striatum (Shi and Cassell, 1998a, b; Price, 2007). Nonetheless, these results do fit with out prior work where inactivation of vlOFC before acquisition in males did not significantly impact safety learning or later recall (Sarlitto et al., 2018).

4.4.2 Insular Cortex

IC is a region known to involved in salience detection, emotional learning and responding to multisensory stimuli (Rodgers et al., 2008). Here we found a main effect of cue in mIC, with increased Fos in the Discrimination condition compared to Controls. Fos in mIC also significantly correlated with both BLA and IL in the Discrimination condition. Multiple regression analysis was performed to see if the combination of BLA, IL, and BLA Fos could predict discrimination index, but no combination of structures reached significance. In the Fear Only condition there was a significant correlation of pIC and BLA Fos, while BLA and aIC Fos significantly correlated in the Control condition. Together, there seems to be an interesting relationship between BLA and subregions of IC along the rostral/caudal axis. IC subregions are highly interconnected and they may interact to mediate BLA activity depending on the level of fear (Shi and Cassell, 1998b). We previously found a necessity of pIC, but not aIC or mIC, in the acquisition of a conditioned inhibition of fear, as measured by a summation test (Foilb et al., 2016). mIC, but not pIC, activation is significant in the Discrimination condition, as is functional connectivity with BLA and IL. This provides evidence

that fear discrimination and conditioned inhibition may arise through distinct mechanisms.

There is substantial evidence in human research and within the PTSD population for IC involvement in fear discrimination. fMRI data finds significantly greater responding in IC during presentations of A compared to presentations of B in discrimination acquisition (Schiller et al., 2008). IC is also a site of functional and structural abnormalities in anxiety and PTSD (Paulus and Stein, 2006; Hughes and Shin, 2011), including the noted correlation of reduced insula volume with reduced safety learning (Gutman et al., 2010). This evidence, along with the correlation results with BLA throughout IC, suggest that IC may work in concert with the amygdala during fear and safety learning.

4.4.3 Extended Amygdala

Fos in BLA was increased in shocked conditions (Fear Only and Discrimination) compared to Controls, but did not vary by sex or correlate with discrimination. Although there is evidence that some lateral amygdala neurons respond specifically to safe cues, many of these neurons reduce firing to the safe cue (Sangha et al., 2013) and safety learning leads to weakened responses in the BLA from auditory inputs (Rogan et al., 2005). Inhibition of specific neuronal populations within the BLA would not be evident in Fos analyses, and is one possible reason why discrimination-specific effects were not found in the BLA. The BLA is highly connected to other regions that appear to be involved in safety

learning, including the vmPFC, OFC, IC and striatum (Kelley et al., 1982; Shi and Cassell, 1998a, b; Ongür and Price, 2000; Cho et al., 2013). Despite similar levels of BLA Fos in Fear Only and Discrimination conditions, it is possible that these connected regions are modulating safety learning via inhibition of specific BLA neurons, which can not be assessed with the existing Fos data.

Due to its role in fear, increased CeA Fos in Fear Only and Discrimination conditions compared to control animals was an anticipated result (LeDoux et al., 1988; Swanson and Petrovich, 1998; Maren, 2001). CeA was the only brain structure explored where Fos significantly correlated with discrimination behavior. CeA Fos and discrimination index correlated such that animals with greater discrimination between A and B cues had increased CeA Fos. CeA also did not correlate with any other brain structures in the Discrimination condition, despite being the only significant predictor of discrimination. This lack of CeA functional connectivity in the Discrimination condition with any other brain regions explored in this study suggests that there is a component missing from the current data—cell type or temporal specificity may be needed or CeA may receive inputs from a structure not included in the current analysis.

The role of CeA in discrimination learning could also be related to the region's function in reward-related behaviors (Kim et al., 2017), since safety signals may have reward-like connotations. Lesions of the CeA, but not the BLA, impair appetitive associative learning (Gallagher et al., 1990; Parkinson et al.,

2008). An appetitive Pavlovian conditioning experiment by Parkinson and colleagues (2008) found that lesions of CeA reduced reward seeking after a cue that signaled reward (a CS+), but did not impact low responding to a distinct cue that indicated the absence of reward (a CS-). As a result of low responding to the CS+, discrimination between cues was disrupted. Further, the role of CeA in reward-related behaviors appears to be mediated by inhibitory projections from the CeA to the vmPFC (Seo et al., 2016). While we did not find evidence for functional connectivity between CeA and vmPFC in the work presented here, this connection may only be present in CeA GABAergic neurons, which were not identified in this study.

While discrimination learning was vastly different in males and females, CeA correlated with discrimination behavior in males and females similarly. That no sex difference was found in any correlation with discrimination was surprising and could potentially indicate that different populations of neurons may be activated in males and females rather than different numbers of neurons. CeA is known to contain receptors for estrogen, progesterone, and androgen, which may be capable of modulating emotional behavior (Toufexis, 2007). Identifying the activation of neurons with these sex-related hormone receptors may be informative in determining if CeA activation is impacting discrimination in sex-specific ways.

The CeA is also comprised of two distinct nuclei – the centromedial (CEm) and centrolateral amygdala (CEl) – that are interconnected but receive different

inputs and consist of different cell types that uniquely alter circuit functioning. Fos analysis here included primarily CEl, although did not carefully exclude CEm. In CEl, neurons expressing protein kinase C-delta inhibit output neurons of the CEm, while neurons expressing neuropeptide somatostatin in the CEl inhibit the paraventricular nucleus of the thalamus and periaqueductal gray, which are both regions responsible for fear expression (for review, see Keifer et al., 2015). With Fos analysis focused on CEl and a significant correlation between Fos and discrimination index, such that increased CeA Fos corresponds with a lower discrimination index indicative of greater discrimination, there may be increased activation within inhibitory neurons in the CEl, which allow for a reduction of fear expression in response to presentation of the safety cue.

Using an explicitly unpaired CS and footshock US, Amano et al. (2010) found that BLA evoked inhibition of CEm was present in both unpaired conditioned animals and animals that underwent fear extinction, but each was through a distinct population of inhibitory neurons. BLA evoked responses in CEl were more responsive in the unpaired condition compared to fear conditioned, fear extinguished, or naïve animals. This work provides evidence that neuronal activation and inhibition may occur within the CeA through distinct cell types, which cannot be parsed apart with Fos, and provide an intriguing pathway for future investigation. While the analysis here focused on CEl, CEm was not explicitly separated or excluded from CeA Fos analysis. The findings in CeA would be better substantiated with specific analysis of these distinct CeA nuclei.

Exploration of neural activation found that anterior medial BNST was the only region with significantly different activation in animals trained in a fear discrimination paradigm compared to animals that underwent fear only conditioning without a safe cue. The activation of BNST in safety learning fits with prior Fos evidence that medial and ventral BNST regions all show increased activation in safety trained animals compared to naïve animals (Campeau et al., 1997). BNST is also anatomically connected to various regions that seem to be involved in safety learning. Anterior medial BNST, the subregion explored for Fos activation, receives projections from IL and is densely connected to amygdala structures, particularly the CeA (Dong et al., 2001; Wood et al., 2018). Anterior medial BNST also projects to nucleus accumbens, a necessary node for future investigation (Dong and Swanson, 2006).

BNST was the only region to show a sex difference in the number of Fos positive cells, which fits with ample evidence that sex-specific modulation of anxiety may depend on different mechanisms within the BNST (for review, see (Toufexis, 2007). This finding also points to neuromodulators that are known to mediate anxiety in the BNST in sex-specific ways. Stress hormone, corticotrophin releasing factor (CRF) activates different circuits in males and females, with particular differences in BNST functional connectivity, and also in relation to estrogen levels (Salvatore et al., 2018). Further, systemically antagonizing serotonin (5-HT) receptor subtype 2C improves discrimination learning (Foilb and

Christianson, 2016), and 5-HT binding the same receptor subtype activates CRF neurons in the BNST, resulting in increased fear and anxiety (Marcinkiewcz et al., 2016). This may be a mechanism that allows for sex specific modulation of fear in discrimination.

With significantly different Fos levels in all conditions and in males and females of the discrimination condition, it was surprising to find that BNST Fos did not correlate with discrimination behavior. There was a trend in females where more BNST Fos corresponded with greater discrimination between safe and danger cues, and this may have reached significance with a larger sample size. BNST also lacked functional connectivity to any of the others regions explored. This is possibly due to the subregion of BNST that was selected as the focus for our initial investigation, or due to the lack of cell type specificity of Fos. Future work will thoroughly investigate additional BNST subregions in the brain tissue of these animals. Anterior lateral BNST, for example, receives projections from the anterior medial BNST, amygdala, vmPFC and IC, making it a likely site of safety integration (Dong et al., 2001; Dong and Swanson, 2006).

4.4.4 Functional Connectivity

The goal of exploring functional connectivity was to determine the connections of the nodes that make up the neural mechanisms underlying safety learning. While this work provided minimal evidence for the complete circuit that allows safety learning to occur, it does provide encouraging areas for future

investigation. The most interesting finding was that very few functional connections were maintained across each behavioral condition, suggesting task specific network activation. PL and IL correlated in all conditions, likely due to their dense anatomical connectivity. All other region correlations in the Discrimination condition – mIC with IL and BLA – were specific to only the Discrimination condition, indicating that these pathways are involved in discrimination learning, but not fear learning. Adding more nodes to the circuitry laid out here should improve inferences that can be made about the functional connectivity involved in safety learning.

4.4.5 Conclusion

With this work, we aimed to test a hypothetical circuitry underlying fear discrimination by exploring Fos activation in regions likely involved in the inhibition of fear by a safety signal and gained substantial information about each of the nodes we investigated. In almost all regions of interest, we found increased Fos in male and female rats that received fear discrimination conditioning compared to control animals. In all but one case (BNST), fear conditioned and discrimination conditioned animals had similar levels of Fos, indicating that more detailed research is necessary to parse apart these two behaviorally similar learning processes. Importantly, none of our findings eliminate the possibility that any brain region explored is part of the circuitry underlying safety learning.

As outlined earlier, there are limitations to using Fos as a neural maker, which means that future investigations are necessary to draw more precise conclusions. Fos doesn't identify 1) cell types: different numbers of excitatory and inhibitory neurons could activated in each condition, 2) neural projections: pathways may be condition-dependent, despite similar Fos activation, 3) temporal specificity: it is unknown which, if any, neurons are responding specifically to presentations the safety cue. For these reasons, regions that displayed similar Fos activation in Fear Only and Discrimination conditions, and even regions that are similarly activated in all conditions, may be involved in the mechanisms that underlie discrimination in a way that cannot be observed with only Fos. Another consideration is whether the current approach was appropriately powered to detect relatively small differences that may exist between sexes or between Discrimination and Fear Only conditioning. Although when combining males and females, the discrimination sample size is 14-16 animals, which is substantial compared to similar extant literature.

Fos activation is able to indicate regions that may be particularly critical to, or nuanced in, safety learning. CeA and BNST appear to be particularly interesting sites of future investigation, as CeA was the only region where a correlation with discrimination index was observed and Fos in BNST was the only sex difference uncovered, as well as the only region where discrimination conditioning produced significantly different Fos compared to fear conditioning. Future investigations of cell types, subregions and additional neural structures

with known connectivity will better explain how these regions are influencing the modulation of fear. Importantly, follow up research can take a mechanistic approach to excite or inhibit regions and pathways that appear to be specific for discrimination learning to determine if the functional connectivity accurately represents the underlying circuitry.

CHAPTER 5

Discussion and Future Directions

5.1 Overview

The goals of this dissertation were to uncover the neural circuitry that allows for appropriate discrimination between cues indicating safety and danger, and to determine if male and females learn and use safety cues differently. First, we took a mechanistic approach to test the necessity of ventrolateral orbitofrontal cortex (vlOFC), ventral hippocampus (VH) and insular cortex (IC) in the acquisition and recall of fear discrimination, as well as the role of IC in the more complex process of conditioned fear inhibition (Chapter 2). Since all of the mechanistic studies were performed in males, exploring potential sex differences in fear discrimination was a critical next step. In Chapter 3, we used a large sample size to compare males and females in safety learning and recall, as well as recall of conditioned inhibition of fear. After finding that females exhibit greater fear inhibition to the safety cue than males during acquisition, we aimed to uncover the neural mechanisms that may underlie the behavioral sex difference. Drawing upon the existing literature on safety learning (reviewed in Chapter 1), I developed a hypothetical neural circuit for safety learning, which I then tested empirically by examining neural activation in male and females that underwent discrimination conditioning compared to animals that underwent fear conditioning or no conditioning. In this concluding chapter, I will review our findings on sex differences in safety learning and the neural nodes involved in safety processing, as well as outline the novel questions for future exploration.

5.2 Sex Differences in Safety Learning

In Chapter 3 we found profound sex differences in fear discrimination learning and recall. Using a large sample size of 60 animals of each sex, we conditioned animals with an AX+/BX- fear discrimination conditioning paradigm, as we had done previously with only males (Chen et al., 2016; Foilb and Christianson, 2016; Foilb et al., 2016; Sarlitto et al., 2018). As in much of our prior work, animals of both sexes significantly discriminated between danger (A) and safety (B) cues during conditioning, but females displayed significantly reduced fear to cue B, as well as significantly greater discrimination as measured by a discrimination index. We were also able to replicate this sex difference in acquisition of fear discrimination in the work presented in Chapter 4. In recall testing, this pattern persisted, with females displaying significantly greater discrimination and significantly less fear to the safety cue compared to males. Interestingly, by the end of the recall test, females and males were responding similarly to A and B cues. Additionally, using repeated conditioning and recall testing to establish the B cue as a conditioned fear inhibitor eliminated the behavioral sex differences observed in the single session of fear discrimination conditioning.

Similar patterns of sex differences in discrimination learning have been reported in other labs. Day and colleagues (2016) found that females initially had greater discrimination between A and B cues, but generalized between cues with repeated conditioning, while males continued to discriminate. Day and colleagues

also found that females did not display delayed learning of the B cue as a danger cue, one of the two critical tests of a conditioned fear inhibitor, as defined by Rescorla (1969). We measured conditioned inhibition with a summation test and saw that females were able to reduce fear to the AB compound compared to presentation of the A cue alone. While we did not observe generalization in females, we did find that males continued to improve discrimination across conditioning and testing, whereas females learned to discriminate earlier in conditioning and maintained relatively consistent discrimination throughout sessions. Similar to Day et al., Greiner et al. (*preprint,* 2018) found that females did not reduce freezing behavior to AB compound cues compared to the A cue alone, while males did display conditioned inhibition, with reduced freezing to AB. Interestingly, they found that females significantly altered darting behavior, an active fear response that has been observed in female rats (Gruene et al., 2015), to the compound cue compared to the danger cue alone. Although we looked for darting in the females described in both Chapters 3 and 4, darting occurred too infrequently to be considered in our analyses.

Sex differences in fear discrimination abilities have also been found in humans. In children, trauma history has a larger negative impact on fear discrimination in females than in males (Gamwell et al., 2015). Healthy adult women display less discrimination between safety and danger compared to men, although this difference appears to be primarily due to discrimination deficits in women on hormonal birth control (Lonsdorf et al., 2015). This finding also

indicates a potential impact of estrogen on fear discrimination, which is evident in rodent research as well (Toufexis et al., 2007; Lonsdorf et al., 2015). Despite the different findings across species and paradigms, that females and males use safety cues differently is a common theme. Interestingly, both men and women with PTSD display abnormal fear discrimination, which may stem from an effect of stress exposure on the neural circuitry involved in safety signal processing (Jovanovic et al., 2009, 2012, 2013) and it is possible that the underlying neural circuitry, and therefore the effects of stressors, are sex-specific. For these reasons, to best understand PTSD and to identify better treatments for fear regulation in PTSD, more information is needed concerning the neural basis of safety.

5.3 The Neural Correlates of Safety Learning

In Chapter 2, we used mechanistic experiments to test the necessity of three brain structures hypothesized to play a role in safety learning or recall – vlOFC, VH, and IC. Inhibition of VH before conditioning reduced fear to all cues in later recall, while inhibition of VH before discrimination recall had no effect. From this we concluded that VH plays a role in fear learning, as previously found (Richmond et al., 1999; Bast et al., 2001; Zhang et al., 2001; Esclassan et al., 2009; Czerniawski et al., 2012; Wang et al., 2012; Cox et al., 2013; Zhang et al., 2014), but does not seem specifically necessary for fear discrimination (Chen et al., 2016).

Despite these findings, others have presented evidence for a role of hippocampus in fear discrimination. Opposing our findings, Heldt et al. (2002), found that hippocampal lesions, ranging from dorsal to ventral, prevent recall, but not acquisition of safety learning. In a mouse model, prevention of hippocampal neurogenesis prevents safety learning, providing potential evidence for the role of hippocampus in this task (Pollak et al., 2008). The hippocampus is also implicated in PTSD, particularly in the presence of danger or threat (Fragkaki et al., 2016), which could be related to the results others have found during discrimination. Nonetheless, it seems most likely that the role of the hippocampus in fear discrimination can be reduced to the structure's role in fear processing, rather than safety learning. For that reason, hippocampus was not explored for neural activation in Chapter 4.

Inhibition of vlOFC before fear discrimination recall resulted in increased fear specifically to the safe cue, while vlOFC inhibition before acquisition had no effect on safety learning or later recall (Sarlitto et al., 2018). Since that work was performed in only males, vlOFC was a region of interest for potential sex differences in activation in Chapter 4. The anatomy of vlOFC – with reciprocal connectivity to the amygdala, similar projections to midbrain regions as amygdala, and projections to striatum and insula – situate vlOFC to be a modulator of fear responding (Swanson and Petrovich, 1998; LeDoux, 2000; Price, 2007; Shi and Cassell, 1998a, b). Investigations of vlOFC Fos in Chapter 4 found no effect of discrimination conditioning compared to fear conditioning or

controls, and no sex differences were found in any condition. In the discrimination condition, vlOFC did not show functional connectivity with any other regions explored. This result, along with the results of the mechanistic work, indicates that vlOFC may play a role in discrimination recall, but not acquisition.

Posterior IC (pIC), but not anterior (aIC) or medial (mIC), was necessary for acquisition of conditioned fear inhibition, as measured by a summation test (Foilb et al., 2016). While NMDA blockade of pIC did not impact fear discrimination, it is possible that pIC is engaged, but not critical, in these early phases of safety learning. We hadn't previously tested the roles of aIC or mIC in fear discrimination and the interconnectivity between IC subregions, as well as their connectivity to sensory thalamus, striatum, ventral medial prefrontal cortex (vmPFC), vlOFC, and amygdala, well positions these regions to be involved in fear modulation to a safety signal (McGeorge and Faull, 1989; Shi and Cassell, 1998a, b). IC also has known roles in salience detection and multisensory integration (Rodgers et al., 2008; Gogolla, 2017). These functions in combination with connections to basolateral amygdala (BLA) and central amygdala (CeA) make IC a likely site of convergence for information about danger and safety.

In our Fos exploration of aIC, mIC, and pIC in Chapter 4, we found that only mIC had significantly different Fos levels based on experimental condition, with increased mIC Fos in animals that underwent discrimination conditioning compared to controls. Fos in IC did not significantly correlate with discrimination index, but interestingly, Fos in the BLA displayed functional connectivity with

different subregions of IC in each behavioral condition. In the discrimination condition, BLA and mIC Fos positively correlated signifying an excitatory relationship, and perhaps relevant to the increased mIC Fos in animals that underwent discrimination conditioning. Due to the bidirectional connectivity between BLA and IC and the lack of directional information provided by Fos, the direction of this excitatory functional connection is currently unknown (Shi and Cassell, 1998a, b). In control animals, BLA and aIC Fos correlated, while pIC and BLA Fos correlated in fear conditioned animals. This provides evidence that IC may integrate information about fear and safety in differently along its rostral-caudal axis. This is likely due to the distinct anatomical connectivity in aIC and pIC, with mIC comprising of a mixture of the anterior and posterior subregions in its connectivity (Gogolla, 2017). For example, aIC only weakly projects to amygdala, while mIC and pIC send dense projections to BLA, as well as to CeA (Shi and Cassell, 1998a). Interestingly, in the discrimination condition, mIC Fos significantly correlated with both IL Fos and BLA Fos, however a multiple regression of these three structures with discrimination index was insignificant. This intriguing structure and its functional connections are key sites of future research, where additional Fos observations, temporal specificity and tract specificity, such as pathway specific opto- or chemogenetic manipulations, would provide important insights to the roles of these structures in safety learning.

Human studies also reveal IC as an interesting region in fear discrimination. An fMRI study found increased activation of IC during danger

cues compared to safe cues in a discrimination paradigm (Schiller et al., 2008). In other paradigms, expectations of danger and pain correlate with activation of IC (Ploghaus et al., 1999; Phelps et al., 2001). Individuals with PTSD also show different resting state functional connectivity in IC compared to healthy controls (Zhang et al., 2016). Future work to further the understanding of the role of IC in fear and the modulation of fear will lead to a better understanding of deficits in fear modulation observed in individuals with PTSD, as stress and trauma may disrupt IC functional connectivity, resulting in the promotion of fear expression.

The exploration of neural activation in Chapter 4 went beyond the brain structures described in Chapter 2. Prelimbic (PL) and infralimbic (IL) subregions of the vmPFC were important regions to explore due in part to their roles in the promotion and extinction of fear, respectively (Sotres-Bayon and Quirk, 2010; Sierra-Mercado et al., 2011). We found significantly increased Fos in both PL and IL in animals that had underwent discrimination conditioning and fear conditioning compared to control animals. The activation of PL and IL in discrimination fits with the existing literature implicating PL and IL in fear discrimination, particularly in response to the danger cue (Sangha et al., 2014). We also found a trending correlation between Fos in PL and discrimination index, where more PL Fos corresponded with greater discrimination. This somewhat contradicts mechanistic studies where inhibition of either PL or IL in a fear discrimination paradigm reduces fear to danger (Sangha et al., 2014). This indicates that Fos observed in PL could represent inhibitory neurons, although it also possible that some PL

activation is specifically in response to accurate discrimination, which fits with work showing increased synchrony between firing in the vmPFC with firing in the BLA with learned discrimination (Likhtik et al., 2014; Stujenske et al., 2014). While we did not observe functional connectivity between BLA and vmPFC, this could be due to the limited number of observations or the lack of temporal specificity provided by Fos.

In the BLA, we found increased Fos in animals that underwent conditioning for discrimination and fear alone compared to control animals. However, BLA Fos did not predict discrimination abilities, despite substantial evidence for the role of BLA in discrimination and safety learning. Amygdala neurons fire during both safety and danger signals and some neurons in the lateral amygdala fire exclusively to cues indicating safety (Genud-Gabai et al., 2013). Yet, it is worth noting that many neurons also inhibit firing to safety signals, which may not be well captured in Fos measurements and could explain the lack of discrimination-specific activation that we observed (Sangha et al., 2013). Safety learning also leads to weakened responses in the BLA from auditory inputs (Rogan et al., 2005), perhaps as a function of decreased amygdala synapse size (Ostroff et al., 2010). As described in reviewing other nodes of this safety learning circuit, BLA is also highly connected to other regions that appear to be involved in safety learning, including the insula, striatum and vlOFC (Kelley et al., 1982; Shi and Cassell, 1998a, b; Ongür and Price, 2000; Cho et al., 2013). Advancing the work of Chapter 4 by looking the projections of

activated BLA neurons to CeA and BNST, as well as to mIC, may better describe the role of BLA in fear discrimination.

CeA and BNST are the major output regions of BLA, and in turn project to the hypothalamus and brainstem regions, which mediate the expression of fear, including freezing, autonomic arousal, hormone release, analgesia, and startle (LeDoux et al., 1988; Van de Kar et al., 1991; Davis, 1992; Kapp BS, 1992). We were surprised to find that Fos in CeA was the only brain region explored that significantly correlated with discrimination index, particularly since increased CeA Fos signified greater discrimination.

BNST has been previously implicated in safety learning with increased activation to a conditioned fear inhibitor than to a fear CS (Campeau et al., 1997) and decreased activation to presentation of a safety signal produced through backwards conditioning (Christianson et al., 2011). In Chapter 4, we found increased activation of the anterior medial subdivision of the BNST after discrimination conditioning compared to both fear conditioned and control animals. BNST was also the only brain region explored in Chapter 4 where we found a sex difference, with less activation in females compared to males. Finding this sex difference in BNST is unsurprising since BNST is well established as a sexually dimorphic region, with larger volumes in males in both rodents and humans (for review, see Goode and Maren, 2017). Evidence for sex differences in stress and anxiety have also been reported in the BNST (Toufexis, 2007; Salvatore et al., 2018).

BNST Fos did not correlate to discrimination index, although there was a trend in females where more BNST Fos corresponded with greater discrimination, which is the same direction as was found with CeA Fos. Anterior medial BNST and CeA are bidirectionally connected, which is one of many pathways that BNST may impact fear discrimination through (Dong et al., 2001; Wood et al., 2018). BNST receives projections from IL and the nucleus accumbens (NAcc; Dong and Swanson, 2006; Wood et al., 2018). Other subregions of BNST are also important areas for future investigation, such as the anterior lateral, which receives projections from anterior medial BNST, amygdala, vmPFC and IC, positioning this subregion to be a site of safety integration (Dong et al., 2001; Dong and Swanson, 2006).

One region that was part of the hypothetical safety learning circuitry outlined in Chapter 1 that was not further explored in Chapter 4 is striatum. The anatomical connectivity of the striatum also places the structure to receive information about proximal danger or safety. Striatum receives projections from the vlOFC and IC (Price, 2007; McGeorge & Faull, 1989) and it is also reciprocally connected with the BLA (Cho et al., 2013; Kelley et al., 1982). In human fMRI research striatum displays increased activation during presentations of danger cues compared to safety cues (Schiller et al., 2008). In mice, safety learning led to a strengthening of responses in the caudate putamen (Rogan et al., 2005) and accidental lesion of the striatum in non-human primates is related to an inability to discriminate between danger and safety (Kazama et al., 2012).

However, there is evidence that NAcc is not necessary for safety learning (Josselyn et al., 2005), thus, future analyses will seek to better characterize the striatal subregions that contribute to safety learning.

It is important to note that there are a number of caveats to interpreting Fos data. For one, null findings in Fos cannot simply be interpreted as a lack of neuronal activation. Activation of brain regions have been observed without changes in Fos protein, and in fact, Fos counts can be inconsistent with measurements of *c-fos* mRNA (for review see, McReynolds et al., 2018). As previously described, Fos also lacks specificity in many ways. Fos is temporally imprecise and cannot be used to assess instantaneous changes in neurons during behavior. Fos is also expressed across various cell types, such that Fos may represent inhibitory or excitatory neurons. Pathway information is also not provided by the Fos studies presented in this dissertation; the projections and inputs of the activated neurons in regions of interest remain unknown. With these limitations in mind, the findings from these studies do allow to us to support, as well as question, the hypothetical circuitry for safety learning presented in Chapters 1 and 4.

5.4 Clinical Implications

It is well established that individuals with PTSD display deficits in appropriate fear modulation to cues indicating safety, while healthy individuals reliably display appropriate safe/danger discrimination in the same paradigm

(Jovanovic et al., 2005, 2009, 2010; Jovanovic and Norrholm, 2011; Jovanovic et al., 2012, 2013). A useful start for better understanding the abnormalities in fear discrimination in individuals with PTSD would be a brain imaging study to look at neural activation during fear discrimination in healthy individuals and PTSD populations. Currently, all evidence for neural bases of abnormal discrimination in individuals with PTSD stems from indirect comparisons of discrimination correlates in healthy individuals and regions known to be differently activated in individuals with PTSD.

Many of the results in neural activation reported in Chapter 4 do correspond with prior work in individuals with PTSD. When compared to healthy controls, individuals with PTSD display structural differences in vmPFC and differences in resting state functional connectivity in IC (Corbo et al., 2005; Etkin and Wagner, 2007; Hughes and Shin 2011; Zhang et al., 2016). IC volume has also been found to correlate with discrimination in individuals with PTSD; such that poor fear inhibitors had smaller IC volume (Gutman et al., 2010). The neural activation and functional connectivity of PL, IL and IC subregions in discrimination learning found in our Fos investigation indicate that abnormalities in these structures could play a role in the deficits in safety learning seen in individuals with PTSD. Significant work still needs to be done in order to translate what we know about the neural mechanisms of safety in animals to the human population.

5.5 Neuromodulators

We have begun to discover ways that safety learning can be improved through pharmacological manipulations. Further investigation of pharmacological interventions that recuperate safety learning deficits would significantly improve the lives of individuals with PTSD who are suffering from debilitating fear. There is evidence that neurotransmitters including dopamine and serotonin impact fear discrimination, making them potential targets of future treatments. Lesions to the serotonergic dorsal raphe nucleus (DRN) impair differential learning to a partially reinforced safety signal and work from our lab found that systemically antagonizing serotonin receptor subtype 2C before AX+/BX- discrimination conditioning improved fear discrimination (Berg et al., 2014; Foilb and Christianson, 2016). Receptors for dopamine have been found important for appropriate modulation of responding in uncertainty paradigms (Larkin et al., 2016) and systemically impacting dopamine D1 receptors impairs fear suppression to a safety cue, with similar results observed when acting on D1 receptors in the BLA (Ng et al., 2018). Further, sex differences have been observed in the serotonergic and dopamine systems, making these neurotransmitters a potential source of the behavioral sex difference reported in Chapter 3 (Mitsushima et al., 2006; Goel and Bale, 2010).

Hormones related to stress, such as corticotrophin releasing factor (CRF) and corticosterone (CORT) are also prospective modulators of fear discrimination and the observed sex differences in safety learning. Sex differences have been

reported in stress reactivity after fear conditioning, such that females display greater activation of the HPA axis, resulting in higher levels of CRF and CORT, despite displaying lower levels of fear compared to males (Oyola and Handa, 2017; Daviu et al., 2014). Sex differences in functional connectivity were observed after CRF administration, particularly in functional connectivity to BNST, which is relevant to our Fos findings in Chapter 4. Interactions of CRF and estrogen levels were also observed in females (Salvatore et al., 2018). Varying levels of these hormones during safety learning in males and females could potentially contribute to the sex differences found in Chapter 3. Further, evidence for hormone modulation in fear discrimination learning provides potential methods for treating those with safety learning impairments.

5.5 Conclusions

Understanding the neural circuit for safety learning will inform development of more effective treatments for anxiety and PTSD that would be profoundly beneficial to those individuals that suffer from debilitating fear. The work of this dissertation has provided further support for some nodes and pathways of the circuitry that I hypothesized underlies safety learning (presented again here for reference, Figure 5.1). While the limitations of Fos do not allow us to eliminate any regions from this circuit, it does provide important objectives of future investigation.

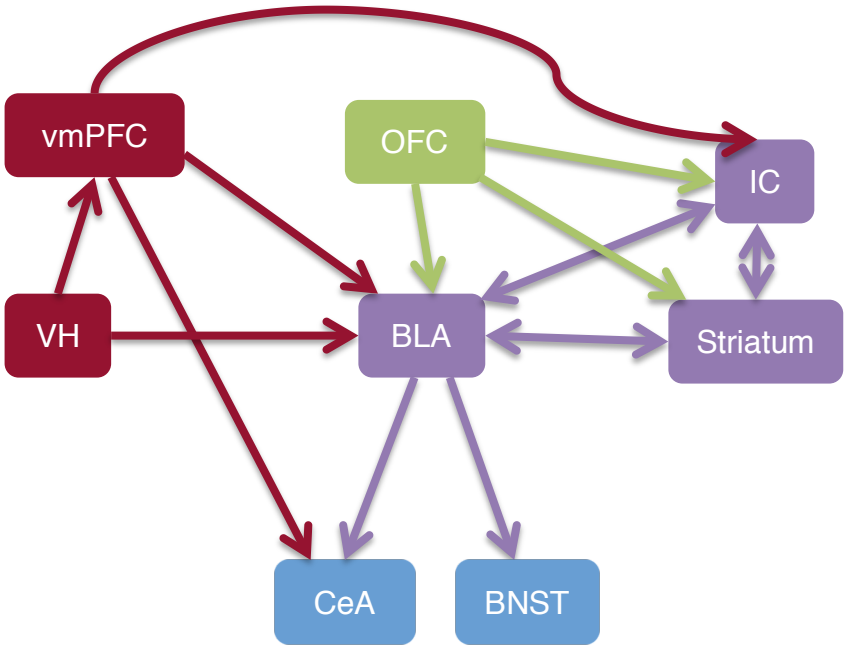

Figure 5.1 A hypothetical circuit for the processing of safety information. As described in Chapters 1 and 4.

While we did not find significantly increased functional connectivity of our regions of interest in fear discrimination conditioned animals compared to fear conditioned or control animals, we did find evidence that select nodes and pathways of interest are likely involved in safety learning. We found functional connectivity between PL and IL subregions of vmPFC and increased PL and IL Fos in discrimination and fear conditioned animals compared to controls. IL Fos also positively correlated with Fos in mIC, indicating that the originally proposed vmPFC to IC pathway may be an excitatory IL projection to mIC. In mIC, there was also increased Fos in discrimination animals compared to controls, providing

further evidence that this region may play a role in safety learning. BLA Fos also correlated with mIC, as proposed in the original hypothetical circuit, although the direction of this excitatory relationship is still unknown.

Despite evidence for functional connectivity of IL and BLA with mIC, findings in the CeA and BNST are perhaps the most fascinating and the best structures for future investigation. CeA was the only brain region analyzed where Fos significantly correlated with discrimination index, however no other region significantly correlated with CeA. This indicates that activation in CeA is somehow modulating fear expression during discrimination learning, but the pathways that allow CeA to do this are undetermined. Tract specific studies looking at direct CeA inputs from BLA, IC, IL, or thalamus would be informative in understanding how CeA might mediate discrimination learning (Keifer et al., 2015).

Also responsible for fear expression, Fos analysis revealed BNST as the only explored region where there was a sex difference in the discrimination condition, as well as the only structure where Fos counts were significantly different in the discrimination condition compared to fear conditioned animals. Unlike in CeA, BNST Fos did not correlate with discrimination index, although the relationship was trending in females. It is possible that more observations would make this correlation significant. There were interesting sex differences in functional connectivity that were not discussed in Chapter 4. Males in the discrimination condition had a significant correlation between BNST Fos and BLA

Fos, which is predicted in the hypothesized circuitry, since BLA is known to project to BNST, which in turn projects to structures responsible for the expression of fear (Walker and Davis, 2008). This correlation is surprisingly not significant in females, or in discrimination animals overall. In females that underwent discrimination conditioning, BNST Fos significantly correlated with PL Fos. Remarkably, the region of BNST analyzed, anterior medial BNST, shares only sparse anatomical connectivity with PL, however, anterior medial BNST is interconnected with anterior lateral BNST which has anatomical connectivity to medial prefrontal cortex (Dong et al., 2001; Dong and Swanson, 2006). For this reason, additional subregions of BNST, including anterior lateral, are ongoing regions for Fos analysis in the present study.

Most regions we explored for Fos in Chapter 2 exhibited equal Fos in Fear Only and Discrimination conditions. To establish how, or if, these regions are specifically involved in discrimination will require mechanistic, cell type specific or tract specific follow up studies. While no main effects or significant correlations were found in the discrimination condition for vlOFC, aIC and pIC, mechanistic studies would be useful in definitively ruling out their role in safety learning in both males and females. To best progress the field towards an understanding of the neural mechanisms of safety learning, additional exploration of Fos in striatum, as well as BNST subregions, and follow-up studies on CeA, BNST, and mIC connectivity with BLA and IL seem to be the most promising routes of discovery.

The work presented here, and studies from other labs over the past 10 years, has increased the understanding of safety learning and the neural mechanisms that underlie it. The work of this dissertation allows the field to continue editing the framework for the neural circuit of safety processing. Across species and techniques, particular regions consistently certify their role in this operation. With the entirety of the safety learning literature and findings in mind, we can continue to develop new questions with the new tools and techniques that are advancing the field of neuroscience.

REFERENCES

Amano T, Unal CT, Paré D (2010) Synaptic correlates of fear extinction in the amygdala. Nat Neurosci, 13:489-494.

Amat J, Baratta MV, Paul E, Bland ST, Watkins LR, Maier SF (2005) Medial prefrontal cortex determines how stressor controllability affects behavior and dorsal raphe nucleus. Nat Neurosci, 8:365–371.

Amat J, Christianson JP, Aleksejev RM, Kim J, Richeson KR, Watkins LR, Maier SF (2014) Control over a stressor involves the posterior dorsal striatum and the act/outcome circuit. Eur J Neurosci, 40:2352–2358.

Anagnostaras SG, Gale GD, Fanselow MS (2001) Hippocampus and contextual fear conditioning: recent controversies and advances. Hippocampus, 11:8–17.

Anderson LC, Petrovich GD (2017) Sex specific recruitment of a medial prefrontal cortex-hippocampal-thalamic system during context-dependent renewal of responding to food cues in rats. Neurobiol Learn Mem, 139:11–21.

Baran SE, Armstrong CE, Niren DC, Conrad CD (2010) Prefrontal cortex lesions and sex differences in fear extinction and perseveration. Learn Mem, 17:267–278.

Baratta MV et al. (2016) Stress Enables Reinforcement-Elicited Serotonergic Consolidation of Fear Memory. Biol Psychiatry, 79:814–822.

Baratta MV, Zarza CM, Gomez DM, Campeau S, Watkins LR, Maier SF (2009) Selective activation of dorsal raphe nucleus-projecting neurons in the ventral medial prefrontal cortex by controllable stress. Eur J Neurosci, 30:1111–1116.

Bast T, da Silva BM, Morris RG (2005) Distinct contributions of hippocampal NMDA and AMPA receptors to encoding and retrieval of one-trial place memory. J Neurosci, 25:5845–5856.

Bast T, Zhang WN, Feldon J (2001) The ventral hippocampus and fear conditioning in rats. Different anterograde amnesias of fear after tetrodotoxin inactivation and infusion of the GABA(A) agonist muscimol. Exp Brain Res, 139:39–52.

Benison AM, Rector DM, Barth DS (2007) Hemispheric mapping of secondary somatosensory cortex in the rat. J Neurophysiol, 97:200–207.

Berg BA, Schoenbaum G, McDannald MA (2014) The dorsal raphe nucleus is integral to negative prediction errors in Pavlovian fear. Eur J Neurosci, 40:3096–3101.

Burger T (2018) Gentle introduction to the statistical foundations of false discovery rate in quantitative proteomics. J Proteome Res, 17(1):12-22.

Bermudez-Rattoni F (2014) The forgotten insular cortex: its role on recognition memory formation. Neurobiol Learn Mem, 109:207–216.

Berridge KC, Robinson TE (1998) What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience. Brain Res Brain Res Rev, 28:309–369.

Bouton ME, King DA (1983) Contextual control of the extinction of conditioned fear: tests for the associative value of the context. J Exp Psychol Anim Behav Process, 9:248–265.

Brown JS, Kalish HI, Farber IE (1951) Conditional fear as revealed by magnitude of startle response to an auditory stimulus. Journal of Experimental Psychology, 41:317–328.

Campbell BM, Merchant KM (2003) Serotonin 2C receptors within the basolateral amygdala induce acute fear-like responses in an open-field environment. Brain Res, 993:1–9.

Campeau S, Falls WA, Cullinan WE, Helmreich DL, Davis M, Watson SJ (1997) Elicitation and reduction of fear: behavioural and neuroendocrine indices and brain induction of the immediate-early gene c-fos. Neuroscience, 78:1087–1104.

Campolattaro MM, Freeman JH (2006) Perirhinal cortex lesions impair feature-negative discrimination. Neurobiol Learn Mem, 86:205–213.

Casanova JP, Madrid C, Contreras M, Rodríguez M, Vasquez M, Torrealba F (2016) A role for the interoceptive insular cortex in the consolidation of learned fear. Behav Brain Res, 296:70–77.

Castro DC, Cole SL, Berridge KC (2015) Lateral hypothalamus, nucleus accumbens, and ventral pallidum roles in eating and hunger: interactions between homeostatic and reward circuitry. Front Syst Neurosci, 9:90.

Chen S et al. (2006) Gray matter density reduction in the insula in fire survivors with posttraumatic stress disorder: a voxel-based morphometric study. Psychiatry Res, 146:65–72.

Chen VM, Foilb AR, Christianson JP (2016) Inactivation of ventral hippocampus interfered with cued-fear acquisition but did not influence later recall or discrimination. Behav Brain Res, 296:249–253.

Cho YT, Ernst M, Fudge JL (2013) Cortico-amygdala-striatal circuits are organized as hierarchical subsystems through the primate amygdala. J Neurosci, 33:14017–14030.

Christianson JP et al. (2008) The sensory insular cortex mediates the stress-buffering effects of safety signals but not behavioral control. J Neurosci, 28:13703–13711.

Christianson JP, Fernando AB, Kazama AM, Jovanovic T, Ostroff LE, Sangha S (2012) Inhibition of fear by learned safety signals: a mini-symposium review. J Neurosci, 32:14118–14124.

Christianson JP, Flyer-Adams JG, Drugan RC, Amat J, Daut RA, Foilb AR, Watkins LR, Maier SF (2014) Learned stressor resistance requires extracellular signal-regulated kinase in the prefrontal cortex. Front Behav Neurosci, 8:348.

Christianson JP, Jennings JH, Ragole T, Flyer JG, Benison AM, Barth DS, Watkins LR, Maier SF (2011) Safety signals mitigate the consequences of uncontrollable stress via a circuit involving the sensory insular cortex and bed nucleus of the stria terminalis. Biol Psychiatry, 70:458–464.

Christianson JP et al. (2010) 5-hydroxytryptamine 2C receptors in the basolateral amygdala are involved in the expression of anxiety after uncontrollable traumatic stress. Biol Psychiatry, 67:339–345.

Clark L, Bechara A, Damasio H, Aitken MR, Sahakian BJ, Robbins TW (2008) Differential effects of insular and ventromedial prefrontal cortex lesions on risky decision-making. Brain, 131:1311–1322.

Cora MC, Kooistra L, Travlos G (2015) Vaginal Cytology of the Laboratory Rat and Mouse: Review and Criteria for the Staging of the Estrous Cycle Using Stained Vaginal Smears. Toxicol Pathol, 43:776–793.

Corbo V, Clément MH, Armony JL, Pruessner JC, Brunet A (2005) Size versus shape differences: contrasting voxel-based and volumetric analyses of the

anterior cingulate cortex in individuals with acute posttraumatic stress disorder. Biol Psychiatry, 58:119–124.

Costanzo M, Jovanovic T, Norrholm SD, Ndiongue R, Reinhardt B, Roy MJ (2016) Psychophysiological Investigation of Combat Veterans with Subthreshold Post-traumatic Stress Disorder Symptoms. Mil Med, 181:793–802.

Cousens G, Otto T (1998) Both pre- and posttraining excitotoxic lesions of the basolateral amygdala abolish the expression of olfactory and contextual fear conditioning. Behav Neurosci, 112:1092–1103.

Cox D, Czerniawski J, Ree F, Otto T (2013) Time course of dorsal and ventral hippocampal involvement in the expression of trace fear conditioning. Neurobiol Learn Mem, 106:316–323.

Czerniawski J, Ree F, Chia C, Otto T (2012) Dorsal versus ventral hippocampal contributions to trace and contextual conditioning: differential effects of regionally selective NMDA receptor antagonism on acquisition and expression. Hippocampus, 22:1528–1539.

Dalley JW, Cardinal RN, Robbins TW (2004) Prefrontal executive and cognitive functions in rodents: neural and neurochemical substrates. Neurosci Biobehav Rev, 28:771–784.

Davis M (1992) The role of the amygdala in fear-potentiated startle: implications for animal models of anxiety. Trends Pharmacol Sci, 13:35–41.

Davis M, Walker DL, Miles L, Grillon C (2010) Phasic vs sustained fear in rats and humans: role of the extended amygdala in fear vs anxiety. Neuropsychopharmacology, 35:105–135.

Daviu N, Andero R, Armario A, Nadal R (2014) Sex differences in the behavioural and hypothalamic-pituitary-adrenal response to contextual fear conditioning in rats. Horm Behav, 66:713–723.

Day HL, Reed MM, Stevenson CW (2016) Sex differences in discriminating between cues predicting threat and safety. Neurobiol Learn Mem, 133:196–203.

de Oliveira AR, Reimer AE, de Macedo CE, de Carvalho MC, Silva MA, Brandao ML (2011) Conditioned fear is modulated by D2 receptor pathway connecting the ventral tegmental area and basolateral amygdala. Neurobiol Learn Mem, 95:37–45.

Do-Monte FH, Quiñones-Laracuente K, Quirk GJ (2015) A temporal shift in the circuits mediating retrieval of fear memory. Nature, 519:460–463.

Dong HW, Petrovich GD, Swanson LW (2001) Topography of projections from amygdala to bed nuclei of the stria terminalis. Brain Res Brain Res Rev, 38:192–246.

Dong HW, Swanson LW (2006) Projections from bed nuclei of the stria terminalis, anteromedial area: cerebral hemisphere integration of neuroendocrine, autonomic, and behavioral aspects of energy balance. J Comp Neurol, 494:142–178.

Dragunow M, Faull R (1989) The use of c-fos as a metabolic marker in neuronal pathway tracing. J Neurosci Methods, 29:261–265.

Engman J, Linnman C, Van Dijk KR, Milad MR (2016) Amygdala subnuclei resting-state functional connectivity sex and estrogen differences. Psychoneuroendocrinology, 63:34–42.

Esclassan F, Coutureau E, Di Scala G, Marchand AR (2009) Differential contribution of dorsal and ventral hippocampus to trace and delay fear conditioning. Hippocampus, 19:33–44.

Estes WK, Skinner BF (1941) Some quantitative properties of anxiety. Journal of Experimental Psychology, 29:390–400.

Etkin A, Wager TD (2007) Functional neuroimaging of anxiety: a meta-analysis of emotional processing in PTSD, social anxiety disorder, and specific phobia. Am J Psychiatry, 164:1476–1488.

Falls WA, Davis M (1995) Lesions of the central nucleus of the amygdala block conditioned excitation, but not conditioned inhibition of fear as measured with the fear-potentiated startle effect. Behav Neurosci, 109:379–387.

Falls WA, Davis M (1997) Inhibition of fear-potentiated startle can be detected after the offset of a feature trained in a serial feature-negative discrimination. J Exp Psychol Anim Behav Process, 23:3–14.

Fanselow MS (1980) Conditioned and unconditional components of post-shock freezing. Pavlov J Biol Sci, 15:177–182.

Fanselow MS (1984) Shock-induced analgesia on the formalin test: effects of shock severity, naloxone, hypophysectomy, and associative variables. Behav Neurosci, 98:79–95.

Fanselow MS, Bolles RC (1979) Naloxone and shock-elicited freezing in the rat. J Comp Physiol Psychol, 93:736–744.

Farrell MR, Gruene TM, Shansky RM (2015) The influence of stress and gonadal hormones on neuronal structure and function. Horm Behav, 76:118–124.

Farrell MR, Holland FH, Shansky RM, Brenhouse HC (2016) Sex-specific effects of early life stress on social interaction and prefrontal cortex dendritic morphology in young rats. Behav Brain Res, 310:119–125.

Fenton GE, Halliday DM, Mason R, Bredy TW, Stevenson CW (2016) Sex differences in learned fear expression and extinction involve altered gamma oscillations in medial prefrontal cortex. Neurobiol Learn Mem, 135:66–72.

Fenton GE, Pollard AK, Halliday DM, Mason R, Bredy TW, Stevenson CW (2014) Persistent prelimbic cortex activity contributes to enhanced learned fear expression in females. Learn Mem, 21:55–60.

Foilb AR, Bals J, Sarlitto MC, Christianson JP (2018) Sex differences in fear discrimination do not manifest as differences in conditioned inhibition. Learn Mem, 25:49–53.

Foilb AR, Christianson JP (2016) Serotonin 2C receptor antagonist improves fear discrimination and subsequent safety signal recall. Prog Neuropsychopharmacol Biol Psychiatry, 65:78–84.

Foilb AR, Christianson JP (2018) Brain Mechanisms for Learning and Using Safety Signals. In: Neurobiology of Abnormal Emotion and Motivated Behaviors (Sangha S, Foti D, eds), pp 204–222.

Foilb AR, Flyer-Adams JG, Maier SF, Christianson JP (2016) Posterior insular cortex is necessary for conditioned inhibition of fear. Neurobiol Learn Mem, 134 Pt B:317–327.

Fragkaki I, Thomaes K, Sijbrandij M (2016) Posttraumatic stress disorder under ongoing threat: a review of neurobiological and neuroendocrine findings. Eur J Psychotraumatol, 7:30915.

Gallagher M, Graham PW, Holland PC (1990) The amygdala central nucleus and appetitive Pavlovian conditioning: lesions impair one class of conditioned behavior. J Neurosci,10:1906–1911.

Gallagher M, McMahan RW, Schoenbaum G (1999) Orbitofrontal cortex and representation of incentive value in associative learning. J Neurosci, 19:6610–6614.

Gallistel CR (1989) Animal cognition: the representation of space, time and number. Annu Rev Psychol, 40:155–189.

Gamwell K, Nylocks M, Cross D, Bradley B, Norrholm SD, Jovanovic T (2015) Fear conditioned responses and PTSD symptoms in children: Sex differences in fear-related symptoms. Dev Psychobiol, 57:799–808.

Gastelum ED, Guilhardi P, Burwell RD (2012) The effects of combined perirhinal and postrhinal damage on complex discrimination tasks. Hippocampus, 22:2059–2067.

Genud-Gabai R, Klavir O, Paz R (2013) Safety signals in the primate amygdala. J Neurosci, 33:17986–17994.

Gewirtz JC, Falls WA, Davis M (1997) Normal conditioned inhibition and extinction of freezing and fear-potentiated startle following electrolytic lesions of medical prefrontal cortex in rats. Behav Neurosci, 111:712–726.

Goel N, Bale TL (2010) Sex differences in the serotonergic influence on the hypothalamic-pituitary-adrenal stress axis. Endocrinology, 151:1784–1794.

Gogolla N (2017) The insular cortex. Curr Biol, 27:R580–R586.

Gogolla N, Takesian AE, Feng G, Fagiolini M, Hensch TK (2014) Sensory integration in mouse insular cortex reflects GABA circuit maturation. Neuron, 83:894–905.

Goode TD, Maren S (2017) Role of the bed nucleus of the stria terminalis in aversive learning and memory. Learn Mem, 24:480–491.

Greenwood BN, Strong PV, Brooks L, Fleshner M (2008) Anxiety-like behaviors produced by acute fluoxetine administration in male Fischer 344 rats are prevented by prior exercise. Psychopharmacology (Berl), 199:209–222.

Greiner EM, Norris MR, M√ºller I, Ng KH, Sangha S (2018) Sex differences in fear regulation and reward seeking behaviors in a fear-safety-reward discrimination task. bioRxiv, 390377.

Gremel CM, Costa RM (2013) Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. Nat Commun, 4:2264.

Gruene TM, Flick K, Stefano A, Shea SD, Shansky RM (2015a) Sexually divergent expression of active and passive conditioned fear responses in rats. Elife, 4

Gruene TM, Roberts E, Thomas V, Ronzio A, Shansky RM (2015b) Sex-specific neuroanatomical correlates of fear expression in prefrontal-amygdala circuits. Biol Psychiatry, 78:186–193.

Gutman D et al. (2010) Structural Brain Differences in PTSD are Associated with Impaired Fear Inhibition. Poster at American College of Neuropsychopharmacology,

Hakamata Y et al. (2007) Structure of orbitofrontal cortex and its longitudinal course in cancer-related post-traumatic stress disorder. Neurosci Res, 59:383–389.

Hammond LJ (1968) Retardation of fear acquisition by a previously inhibitory CS. J Comp Physiol Psychol, 66:756–759.

Hammond LJ, Maser J (1970) Forgetting and conditioned suppression: role of a temporal discrimination. J Exp Anal Behav, 13:333–338.

Hart G, Leung BK, Balleine BW (2014) Dorsal and ventral streams: the distinct role of striatal subregions in the acquisition and performance of goal-directed actions. Neurobiol Learn Mem, 108:104–118.

Heldt SA, Coover GD, Falls WA (2002) Posttraining but not pretraining lesions of the hippocampus interfere with feature-negative discrimination of fear-potentiated startle. Hippocampus, 12:774–786.

Heldt SA, Falls WA (1998) Destruction of the auditory thalamus disrupts the production of fear but not the inhibition of fear conditioned to an auditory stimulus. Brain Res, 813:274–282.

Heldt SA, Falls WA (2003) Destruction of the inferior colliculus disrupts the production and inhibition of fear conditioned to an acoustic stimulus. Behav Brain Res, 144:175–185.

Hikind N, Maroun M (2008) Microinfusion of the D1 receptor antagonist, SCH23390 into the IL but not the BLA impairs consolidation of extinction of auditory fear conditioning. Neurobiol Learn Mem, 90:217–222.

Hobin JA, Ji J, Maren S (2006) Ventral hippocampal muscimol disrupts context-specific fear memory retrieval after extinction in rats. Hippocampus, 16:174–182.

Holland PC (1984) Differential effects of reinforcement of an inhibitory feature after serial and simultaneous feature negative discrimination training. J Exp Psychol Anim Behav Process, 10:461–475.

Holland PC (1989) Transfer of negative occasion setting and conditioned inhibition across conditioned and unconditioned stimuli. J Exp Psychol Anim Behav Process, 15:311–328.

Holland PC (2012) Role of amygdala central nucleus in feature negative discriminations. Behav Neurosci, 126:670–680.

Hoover WB, Vertes RP (2007) Anatomical analysis of afferent projections to the medial prefrontal cortex in the rat. Brain Struct Funct, 212:149–179.

Hughes KC, Shin LM (2011) Functional neuroimaging studies of post-traumatic stress disorder. Expert Rev Neurother, 11:275–285.

Hurley KM, Herbert H, Moga MM, Saper CB (1991) Efferent projections of the infralimbic cortex of the rat. J Comp Neurol, 308:249–276.

Izquierdo A (2017) Functional Heterogeneity within Rat Orbitofrontal Cortex in Reward Learning and Decision Making. J Neurosci, 37:10529–10540.

Izquierdo I, Quillfeldt JA, Zanatta MS, Quevedo J, Schaeffer E, Schmitz PK, Medina JH (1997) Sequential role of hippocampus and amygdala, entorhinal cortex and parietal cortex in formation and retrieval of memory for inhibitory avoidance in rats. Eur J Neurosci, 9:786–793.

Jenewein J et al. (2016) Altered Pain Perception and Fear-Learning Deficits in Subjects With Posttraumatic Stress Disorder. J Pain, 17:1325–1333.

Josselyn SA, Falls WA, Gewirtz JC, Pistell P, Davis M (2005) The nucleus accumbens is not critically involved in mediating the effects of a safety signal on behavior. Neuropsychopharmacology, 30:17–26.

Jovanovic T, Kazama A, Bachevalier J, Davis M (2012) Impaired safety signal learning may be a biomarker of PTSD. Neuropharmacology, 62:695–704.

Jovanovic T, Keyes M, Fiallos A, Myers KM, Davis M, Duncan EJ (2005) Fear potentiation and fear inhibition in a human fear-potentiated startle paradigm. Biol Psychiatry, 57:1559–1564.

Jovanovic T, Norrholm SD (2011) Neural mechanisms of impaired fear inhibition in posttraumatic stress disorder. Front Behav Neurosci, 5:44.

Jovanovic T, Norrholm SD, Blanding NQ, Davis M, Duncan E, Bradley B, Ressler KJ (2010) Impaired fear inhibition is a biomarker of PTSD but not depression. Depress Anxiety, 27:244–251.

Jovanovic T, Norrholm SD, Fennell JE, Keyes M, Fiallos AM, Myers KM, Davis M, Duncan EJ (2009) Posttraumatic stress disorder may be associated with impaired fear inhibition: relation to symptom severity. Psychiatry Res, 167:151–160.

Jovanovic T, Sakoman AJ, Kozarić-Kovačić D, Meštrović AH, Duncan EJ, Davis M, Norrholm SD (2013) Acute stress disorder versus chronic posttraumatic stress disorder: inhibition of fear as a function of time since trauma. Depress Anxiety, 30:217–224.

Kahnt T, Tobler PN (2013) Salience signals in the right temporoparietal junction facilitate value-based decisions. J Neurosci, 33:863–869.

Kapp BS, Whalen PJ, Supple WF, Pascoe JP (1992) Amygdaloid contributions to conditioned arousal and sensory information processing. In: *The amygdala: Neurobiological Aspects of Emotion, Memory, and Mental Dysfunction* ed), pp 229–254. New York, NY: Wiley-Liss.

Kazama AM, Heuer E, Davis M, Bachevalier J (2012) Effects of neonatal amygdala lesions on fear learning, conditioned inhibition, and extinction in adult macaques. Behav Neurosci, 126:392–403.

Keifer OP, Hurt RC, Ressler KJ, Marvar PJ (2015) The Physiology of Fear: Reconceptualizing the Role of the Central Amygdala in Fear Learning. Physiology (Bethesda), 30:389–401.

Keiser AA, Turnbull LM, Darian MA, Feldman DE, Song I, Tronson NC (2017) Sex Differences in Context Fear Generalization and Recruitment of Hippocampus and Amygdala during Retrieval. Neuropsychopharmacology, 42:397–407.

Kelley AE, Domesick VB, Nauta WJ (1982) The amygdalostriatal projection in the rat--an anatomical study by anterograde and retrograde tracing methods. Neuroscience, 7:615–630.

Kesner RP, Churchwell JC (2011) An analysis of rat prefrontal cortex in mediating executive function. Neurobiol Learn Mem, 96:417–431.

Kessler RC, Sonnega A, Bromet E, Hughes M, Nelson CB (1995) Posttraumatic stress disorder in the National Comorbidity Survey. Arch Gen Psychiatry, 52:1048–1060.

Kilpatrick DG, Resnick HS, Milanak ME, Miller MW, Keyes KM, Friedman MJ (2013) National estimates of exposure to traumatic events and PTSD prevalence using DSM-IV and DSM-5 criteria. J Trauma Stress, 26:537–547.

Kim JJ, Jung MW (2006) Neural circuits and mechanisms involved in Pavlovian fear conditioning: a critical review. Neurosci Biobehav Rev, 30:188–202.

Kim J, Zhang X, Muralidhar S, LeBlanc SA, Tonegawa S (2017) Basolateral to Central Amygdala Neural Circuits for Appetitive Behaviors. Neuron,93(6):1464-1479.e5.

Kogler L, Gur RC, Derntl B (2015) Sex differences in cognitive regulation of psychosocial achievement stress: brain and behavior. Hum Brain Mapp, 36:1028–1042.

Kong E, Monje FJ, Hirsch J, Pollak DD (2014) Learning not to fear: neural correlates of learned safety. Neuropsychopharmacology, 39:515–527.

Konorski J (1948) Conditioned reflexes and neuron organization. Cambridge University Press: New York, NY.

Konorski J (1967) Integrative activity of the brain.

Lalumiere RT (2014) Optogenetic dissection of amygdala functioning. Front Behav Neurosci, 8:107.

Larkin JD, Jenni NL, Floresco SB (2016) Modulation of risk/reward decision making by dopaminergic transmission within the basolateral amygdala. Psychopharmacology (Berl), 233:121–136.

Lebron-Milad K, Graham BM, Milad MR (2012) Low estradiol levels: a vulnerability factor for the development of posttraumatic stress disorder. Biol Psychiatry, 72:6–7.

Lebron-Milad K, Milad MR (2012) Sex differences, gonadal hormones and the fear extinction network: implications for anxiety disorders. Biol Mood Anxiety Disord, 2:3.

LeDoux J (1996) Emotional networks and motor control: a fearful view. Prog Brain Res, 107:437–446.

LeDoux JE (2000) Emotion circuits in the brain. Annu Rev Neurosci, 23:155–184.

LeDoux JE (2014) Coming to terms with fear. Proc Natl Acad Sci U S A, 111:2871–2878.

LeDoux JE, Cicchetti P, Xagoraris A, Romanski LM (1990) The lateral amygdaloid nucleus: sensory interface of the amygdala in fear conditioning. J Neurosci, 10:1062–1069.

LeDoux JE, Iwata J, Cicchetti P, Reis DJ (1988) Different projections of the central amygdaloid nucleus mediate autonomic and behavioral correlates of conditioned fear. J Neurosci, 8:2517–2529.

Lee JH, Lee S, Kim JH (2016) Amygdala Circuits for Fear Memory: A Key Role for Dopamine Regulation. Neuroscientist,

Liberzon I, Martis B (2006) Neuroimaging studies of emotional responses in PTSD. Ann N Y Acad Sci, 1071:87–109.

Likhtik E, Stujenske JM, Topiwala MA, Harris AZ, Gordon JA (2014) Prefrontal entrainment of amygdala activity signals safety in learned fear and innate anxiety. Nat Neurosci, 17:106–113.

Lister S, Pearce JM, Butcher SP, Collard KJ, Foster GA (1996) Acquisition of conditioned inhibition in rats is impaired by ablation of serotoninergic pathways. Eur J Neurosci, 8:415–423.

Lonsdorf TB et al. (2015) Sex differences in conditioned stimulus discrimination during context-dependent fear learning and its retrieval in humans: the role of

biological sex, contraceptives and menstrual cycle phases. J Psychiatry Neurosci, 40:368–375.

Lopez-Larson MP, Anderson JS, Ferguson MA, Yurgelun-Todd D (2011) Local brain connectivity and associations with gender and age. Dev Cogn Neurosci, 1:187–197.

Lynch J, Cullen PK, Jasnow AM, Riccio DC (2013) Sex differences in the generalization of fear as a function of retention intervals. Learn Mem, 20:628–632.

MacLeod JE, Bucci DJ (2010) Contributions of the subregions of the medial prefrontal cortex to negative occasion setting. Behav Neurosci, 124:321–328.

Marcinkiewcz CA et al. (2016) Serotonin engages an anxiety and fear-promoting circuit in the extended amygdala. Nature, 537:97–101.

Maren S (2001) Neurobiology of Pavlovian fear conditioning. Annu Rev Neurosci, 24:897–931.

Maren S, Aharonov G, Fanselow MS (1996) Retrograde abolition of conditional fear after excitotoxic lesions in the basolateral amygdala of rats: absence of a temporal gradient. Behav Neurosci, 110:718–726.

Maren S, Holt WG (2004) Hippocampus and Pavlovian fear conditioning in rats: muscimol infusions into the ventral, but not dorsal, hippocampus impair the acquisition of conditional freezing to an auditory conditional stimulus. Behav Neurosci, 118:97–110.

Maren S, Phan KL, Liberzon I (2013) The contextual brain: implications for fear conditioning, extinction and psychopathology. Nat Rev Neurosci, 14:417–428.

Martin JR, Ballard TM, Higgins GA (2002) Influence of the 5-HT2C receptor antagonist, SB-242084, in tests of anxiety. Pharmacol Biochem Behav, 71:615–625.

Mascagni F, McDonald AJ, Coleman JR (1993) Corticoamygdaloid and corticocortical projections of the rat temporal cortex: a Phaseolus vulgaris leucoagglutinin study. Neuroscience, 57:697–715.

McDannald MA, Saddoris MP, Gallagher M, Holland PC (2005) Lesions of orbitofrontal cortex impair rats' differential outcome expectancy learning but not conditioned stimulus-potentiated feeding. J Neurosci, 25:4626–4632.

McDonald AJ (1991) Topographical organization of amygdaloid projections to the caudatoputamen, nucleus accumbens, and related striatal-like areas of the rat brain. Neuroscience, 44:15–33.

McDonald AJ (1998) Cortical pathways to the mammalian amygdala. Prog Neurobiol, 55:257–332.

McDonald AJ, Shammah-Lagnado SJ, Shi C, Davis M (1999) Cortical afferents to the extended amygdala. Ann N Y Acad Sci, 877:309–338.

McGeorge AJ, Faull RL (1989) The organization of the projection from the cerebral cortex to the striatum in the rat. Neuroscience, 29:503–537.

McLean CP, Asnaani A, Litz BT, Hofmann SG (2011) Gender differences in anxiety disorders: prevalence, course of illness, comorbidity and burden of illness. J Psychiatr Res, 45:1027–1035.

McNally GP, Westbrook RF (2006) Predicting danger: the nature, consequences, and neural mechanisms of predictive fear learning. Learn Mem, 13:245–253.

McReynolds JR, Christianson JP, Blacktop JM, Mantsch JR (2018) What does the Fos say? Using Fos-based approaches to understand the contribution of stress to substance use disorders. Neurobiol Stress, 9:271–285.

Meyer HC, Bucci DJ (2014) The contribution of medial prefrontal cortical regions to conditioned inhibition. Behav Neurosci, 128:644–653.

Milad MR, Rosenbaum BL, Simon NM (2014) Neuroscience of fear extinction: implications for assessment and treatment of fear-based and anxiety related disorders. Behav Res Ther, 62:17–23.

Miller RR, Hallam SC, Hong JY, Dufore DS (1991) Associative structure of differential inhibition: implications for models of conditioned inhibition. J Exp Psychol Anim Behav Process, 17:141–150.

Mitsushima D, Yamada K, Takase K, Funabashi T, Kimura F (2006) Sex differences in the basolateral amygdala: the extracellular levels of serotonin and dopamine, and their responses to restraint stress in rats. Eur J Neurosci, 24:3245–3254.

Morgan MA, Pfaff DW (2001) Effects of estrogen on activity and fear-related behaviors in mice. Horm Behav, 40:472–482.

Morris RG (2013) NMDA receptors and memory encoding. Neuropharmacology, 74:32–40.

Moscarello JM, LeDoux JE (2013) Active avoidance learning requires prefrontal suppression of amygdala-mediated defensive reactions. J Neurosci, 33:3815–3823.

Mufson EJ, Mesulam MM (1982) Insula of the old world monkey. II: Afferent cortical input and comments on the claustrum. J Comp Neurol, 212:23–37.

Murray EA, Izquierdo A (2007) Orbitofrontal cortex and amygdala contributions to affect and action in primates. Ann N Y Acad Sci, 1121:273–296.

Myers KM, Davis M (2004) AX+, BX- discrimination learning in the fear-potentiated startle paradigm: possible relevance to inhibitory fear learning in extinction. Learn Mem, 11:464–475.

Ng KH, Pollock MW, Urbanczyk PJ, Sangha S (2018) Altering D1 receptor activity in the basolateral amygdala impairs fear suppression during a safety cue. Neurobiol Learn Mem, 147:26–34.

Ogawa M, van der Meer MA, Esber GR, Cerri DH, Stalnaker TA, Schoenbaum G (2013) Risk-responsive orbitofrontal neurons track acquired salience. Neuron, 77:251–258.

Ongür D, Price JL (2000) The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. Cereb Cortex, 10:206–219.

Orsini CA, Kim JH, Knapska E, Maren S (2011) Hippocampal and prefrontal projections to the basal amygdala mediate contextual regulation of fear after extinction. J Neurosci, 31:17269–17277.

Ostlund SB, Balleine BW (2007) Orbitofrontal cortex mediates outcome encoding in Pavlovian but not instrumental conditioning. J Neurosci, 27:4819–4825.

Ostroff LE, Cain CK, Bedont J, Monfils MH, Ledoux JE (2010) Fear and safety learning differentially affect synapse size and dendritic translation in the lateral amygdala. Proc Natl Acad Sci U S A, 107:9418–9423.

Ostroff LE, Cain CK, Jindal N, Dar N, Ledoux JE (2012) Stability of presynaptic vesicle pools and changes in synapse morphology in the amygdala following fear learning in adult rats. J Comp Neurol, 520:295–314.

Oyola MG, Handa RJ (2017) Hypothalamic-pituitary-adrenal and hypothalamic-pituitary-gonadal axes: sex differences in regulation of stress responsivity. Stress, 20:476–494.

Parkinson JA, Robbins TW, Everitt BJ (2000) Dissociable roles of the central and basolateral amygdala in appetitive emotional learning. Eur J Neurosci, 12:405–413.

Paulus MP, Stein MB (2006) An insular view of anxiety. Biol Psychiatry, 60:383–387.

Pavlov IP (1927) Conditioned reflexes: An investigation of the physiological activity of cerebral cortex (Anrep, G. V., Trans.).

Paxinos G, Watson C (2007) The Rat Brain in Stereotaxic Coordinates, Sixth Edition: Hard Cover Edition. Academic Press:

Pettersson R, Hagsäter SM, Eriksson E (2016) Serotonin depletion eliminates sex differences with respect to context-conditioned immobility in rat. Psychopharmacology (Berl), 233:1513–1521.

Phelps EA (2004) Human emotion and memory: interactions of the amygdala and hippocampal complex. Curr Opin Neurobiol, 14:198–202.

Phelps EA, O'Connor KJ, Gatenby JC, Gore JC, Grillon C, Davis M (2001) Activation of the left amygdala to a cognitive representation of fear. Nat Neurosci, 4:437–441.

Ploghaus A, Tracey I, Gati JS, Clare S, Menon RS, Matthews PM, Rawlins JN (1999) Dissociating pain from its anticipation in the human brain. Science, 284:1979–1981.

Pollak DD, Monje FJ, Zuckerman L, Denny CA, Drew MR, Kandel ER (2008) An animal model of a behavioral intervention for depression. Neuron, 60:149–161.

Price JL (2007) Definition of the orbital cortex in relation to specific connections with limbic and visceral structures and other cortical regions. Ann N Y Acad Sci, 1121:54–71.

Pryce CR, Lehmann J, Feldon J (1999) Effect of sex on fear conditioning is similar for context and discrete CS in Wistar, Lewis and Fischer rat strains. Pharmacol Biochem Behav, 64:753–759.

Quirk GJ, Repa C, LeDoux JE (1995) Fear conditioning enhances short-latency auditory responses of lateral amygdala neurons: parallel recordings in the freely behaving rat. Neuron, 15:1029–1039.

Rauch SL, Shin LM, Phelps EA (2006) Neurocircuitry models of posttraumatic stress disorder and extinction: human neuroimaging research--past, present, and future. Biol Psychiatry, 60:376–382.

Rauch SL, Whalen PJ, Shin LM, McInerney SC, Macklin ML, Lasko NB, Orr SP, Pitman RK (2000) Exaggerated amygdala response to masked facial stimuli in posttraumatic stress disorder: a functional MRI study. Biol Psychiatry, 47:769–776.

Ray MH, Hanlon E, McDannald MA (2018) Lateral orbitofrontal cortex partitions mechanisms for fear regulation and alcohol consumption. PLoS One, 13:e0198043.

Rempel-Clower NL (2007) Role of orbitofrontal cortex connections in emotion. Ann N Y Acad Sci, 1121:72–86.

Remple MS, Henry EC, Catania KC (2003) Organization of somatosensory cortex in the laboratory rat (Rattus norvegicus): Evidence for two lateral areas joined at the representation of the teeth. J Comp Neurol, 467:105–118.

Rescorla RA (1969) Conditioned inhibition of fear resulting from negative CS-US contingencies. J Comp Physiol Psychol, 67:504–509.

Rhodes SE, Killcross AS (2007) Lesions of rat infralimbic cortex result in disrupted retardation but normal summation test performance following training on a Pavlovian conditioned inhibition procedure. Eur J Neurosci, 26:2654–2660.

Riceberg JS, Shapiro ML (2012) Reward stability determines the contribution of orbitofrontal cortex to adaptive behavior. J Neurosci, 32:16402–16409.

Riceberg JS, Shapiro ML (2017) Orbitofrontal Cortex Signals Expected Outcomes with Predictive Codes When Stable Contingencies Promote the Integration of Reward History. J Neurosci, 37:2010–2021.

Richmond MA, Yee BK, Pouzet B, Veenman L, Rawlins JN, Feldon J, Bannerman DM (1999) Dissociating context and space within the hippocampus: effects of complete, dorsal, and ventral excitotoxic hippocampal lesions on conditioned freezing and spatial learning. Behav Neurosci, 113:1189–1203.

Robinson CJ, Burton H (1980a) Organization of somatosensory receptive fields in cortical areas 7b, retroinsula, postauditory and granular insula of M. fascicularis. J Comp Neurol, 192:69–92.

Robinson CJ, Burton H (1980b) Somatic submodality distribution within the second somatosensory (SII), 7b, retroinsular, postauditory, and granular insular cortical areas of M. fascicularis. J Comp Neurol, 192:93–108.

Robinson CJ, Burton H (1980c) Somatotopographic organization in the second somatosensory area of M. fascicularis. J Comp Neurol, 192:43–67.

Robinson S, Keene CS, Iaccarino HF, Duan D, Bucci DJ (2011) Involvement of retrosplenial cortex in forming associations between multiple sensory stimuli. Behav Neurosci, 125:578–587.

Rodgers KM, Benison AM, Klein A, Barth DS (2008) Auditory, somatosensory, and multisensory insular cortex in the rat. Cereb Cortex, 18:2941–2951.

Rogan MT, Leon KS, Perez DL, Kandel ER (2005) Distinct neural signatures for safety and danger in the amygdala and striatum of the mouse. Neuron, 46:309–320.

Rogan MT, Stäubli UV, LeDoux JE (1997) Fear conditioning induces associative long-term potentiation in the amygdala. Nature, 390:604–607.

Rogers-Carter MM, Varela JA, Gribbons KB, Pierce AF, McGoey MT, Ritchey M, Christianson JP (2018) Insular cortex mediates approach and avoidance responses to social affective stimuli. Nat Neurosci, 21:404–414.

Romanski LM, Clugnet MC, Bordi F, LeDoux JE (1993) Somatosensory and auditory convergence in the lateral nucleus of the amygdala. Behav Neurosci, 107:444–450.

Romanski LM, LeDoux JE (1993) Information cascade from primary auditory cortex to the amygdala: corticocortical and corticoamygdaloid projections of temporal cortex in the rat. Cereb Cortex, 3:515–532.

Rosen JB, Hitchcock JM, Miserendino MJ, Falls WA, Campeau S, Davis M (1992) Lesions of the perirhinal cortex but not of the frontal, medial prefrontal, visual, or insular cortex block fear-potentiated startle using a visual conditioned stimulus. J Neurosci, 12:4624–4633.

Rosenzweig MR, Bennett EL (1996) Psychobiology of plasticity: effects of training and experience on brain and behavior. Behav Brain Res, 78:57–65.

Rudebeck PH, Murray EA (2014) The orbitofrontal oracle: cortical mechanisms for the prediction and evaluation of specific behavioral outcomes. Neuron, 84:1143–1156.

Salvatore M, Wiersielis KR, Luz S, Waxler DE, Bhatnagar S, Bangasser DA (2018) Sex differences in circuits activated by corticotropin releasing factor in rats. Horm Behav, 97:145–153.

Sangha S, Chadick JZ, Janak PH (2013) Safety encoding in the basal amygdala. J Neurosci, 33:3744–3751.

Sangha S, Robinson PD, Greba Q, Davies DA, Howland JG (2014) Alterations in reward, fear and safety cue discrimination after inactivation of the rat prelimbic and infralimbic cortices. Neuropsychopharmacology, 39:2405–2413.

Sarlitto MC, Foilb AR, Christianson JP (2018) Inactivation of the Ventrolateral Orbitofrontal Cortex Impairs Flexible Use of Safety Signals. Neuroscience, 379:350–358.

Schiller D, Levy I, Niv Y, LeDoux JE, Phelps EA (2008) From fear to safety and back: reversal of fear in the human brain. J Neurosci, 28:11517–11525.

Seo DO, Funderburk SC, Bhatti DL, Motard LE, Newbold D, Girven KS, McCall JG, Krashes M, Sparta DR, Bruchas MR (2016) A GABAergic Projection from the Centromedial Nuclei of the Amygdala to Ventromedial Prefrontal Cortex Modulates Reward Behavior. J Neurosci, 36:10831–10842.

Shansky RM (2015) Sex differences in PTSD resilience and susceptibility: Challenges for animal models of fear learning. Neurobiol Stress, 1:60–65.

Shansky RM, Woolley CS (2016) Considering Sex as a Biological Variable Will Be Valuable for Neuroscience Research. J Neurosci, 36:11817–11822.

Shi C, Davis M (1999) Pain pathways involved in fear conditioning measured with fear-potentiated startle: lesion studies. J Neurosci, 19:420–430.

Shi CJ, Cassell MD (1998a) Cascade projections from somatosensory cortex to the rat basolateral amygdala via the parietal insular cortex. J Comp Neurol, 399:469–491.

Shi CJ, Cassell MD (1998b) Cortical, thalamic, and amygdaloid connections of the anterior and posterior insular cortices. J Comp Neurol, 399:440–468.

Sierra-Mercado D, Padilla-Coreano N, Quirk GJ (2011) Dissociable roles of prelimbic and infralimbic cortices, ventral hippocampus, and basolateral amygdala in the expression and extinction of conditioned fear. Neuropsychopharmacology, 36:529–538.

Sotres-Bayon F, Quirk GJ (2010) Prefrontal control of fear: more than just extinction. Curr Opin Neurobiol, 20:231–235.

Stalnaker TA, Cooch NK, Schoenbaum G (2015) What the orbitofrontal cortex does not do. Nat Neurosci, 18:620–627.

Stefanacci L, Farb CR, Pitkänen A, Go G, LeDoux JE, Amaral DG (1992) Projections from the lateral nucleus to the basal nucleus of the amygdala: a light and electron microscopic PHA-L study in the rat. J Comp Neurol, 323:586–601.

Stevens JS, Hamann S (2012) Sex differences in brain activation to emotional stimuli: a meta-analysis of neuroimaging studies. Neuropsychologia, 50:1578–1593.

Strigo IA, Simmons AN, Matthews SC, Grimes EM, Allard CB, Reinhardt LE, Paulus MP, Stein MB (2010) Neural correlates of altered pain response in women with posttraumatic stress disorder from intimate partner violence. Biol Psychiatry, 68:442–450.

Stujenske JM, Likhtik E, Topiwala MA, Gordon JA (2014) Fear and safety engage competing patterns of theta-gamma coupling in the basolateral amygdala. Neuron, 83:919–933.

Sudakov K, MacLean PD, Reeves A, Marino R (1971) Unit study of exteroceptive inputs to claustrocortex in awake, sitting, squirrel monkey. Brain Res, 28:19–34.

Sul JH, Kim H, Huh N, Lee D, Jung MW (2010) Distinct roles of rodent orbitofrontal and medial prefrontal cortex in decision making. Neuron, 66:449–460.

Swanson LW, Petrovich GD (1998) What is the amygdala. Trends Neurosci, 21:323–331.

Swartzentruber D (1995) MODULATORY MECHANISMS IN PAVLOVIAN CONDITIONING. Animal Learning & Behavior, 23:123–143.

Toufexis D (2007) Region- and sex-specific modulation of anxiety behaviours in the rat. J Neuroendocrinol, 19:461–473.

Toufexis DJ, Myers KM, Bowser ME, Davis M (2007) Estrogen disrupts the inhibition of fear in female rats, possibly through the antagonistic effects of estrogen receptor alpha (ERalpha) and ERbeta. J Neurosci, 27:9729–9735.

Turner BH, Zimmer J (1984) The architecture and some of the interconnections of the rat's amygdala and lateral periallocortex. J Comp Neurol, 227:540–557.

Van de Kar LD, Piechowski RA, Rittenhouse PA, Gray TS (1991) Amygdaloid lesions: differential effect on conditioned stress and immobilization-induced increases in corticosterone and renin secretion. Neuroendocrinology, 54:89–95.

Vieira PA, Corches A, Lovelace JW, Westbrook KB, Mendoza M, Korzus E (2015) Prefrontal NMDA receptors expressed in excitatory neurons control fear discrimination and fear extinction. Neurobiol Learn Mem, 119:52–62.

Waddell J, Heldt S, Falls WA (2003) Posttraining lesion of the superior colliculus interferes with feature-negative discrimination of fear-potentiated startle. Behav Brain Res, 142:115–124.

Wagner AR, Logan FA, Haberlandt K, Price T (1968) Stimulus selection in animal discrimination learning. J Exp Psychol, 76:171–180.

Walf AA, Frye CA (2005) ERbeta-selective estrogen receptor modulators produce antianxiety behavior when administered systemically to ovariectomized rats. Neuropsychopharmacology, 30:1598–1609.

Walf AA, Rhodes ME, Frye CA (2004) Antidepressant effects of ERbeta-selective estrogen receptor modulators in the forced swim test. Pharmacol Biochem Behav, 78:523–529.

Walker DL, Davis M (2008) Role of the extended amygdala in short-duration versus sustained fear: a tribute to Dr. Lennart Heimer. Brain Struct Funct, 213:29–42.

Wang SH, Finnie PS, Hardt O, Nader K (2012) Dorsal hippocampus is necessary for novel learning but sufficient for subsequent similar learning. Hippocampus, 22:2157–2170.

Watkins LR, Wiertelak EP, McGorry M, Martinez J, Schwartz B, Sisk D, Maier SF (1998) Neurocircuitry of conditioned inhibition of analgesia: effects of amygdala, dorsal raphe, ventral medullary, and spinal cord lesions on antianalgesia in the rat. Behav Neurosci, 112:360–378.

Wheeler AL et al. (2013) Identification of a functional connectome for long-term fear memory in mice. PLoS Comput Biol, 9:e1002853.

Wilson RC, Takahashi YK, Schoenbaum G, Niv Y (2014) Orbitofrontal cortex as a cognitive map of task space. Neuron, 81:267–279.

Winslow JT, Noble PL, Davis M (2008) AX+/BX- discrimination learning in the fear-potentiated startle paradigm in monkeys. Learn Mem, 15:63–66.

Wood M, Adil O, Wallace T, Fourman S, Wilson SP, Herman JP, Myers B (2018) Infralimbic prefrontal cortex structural and functional connectivity with the limbic forebrain: a combined viral genetic and optogenetic analysis. Brain Struct Funct,

Young JJ, Shapiro ML (2011) The orbitofrontal cortex and response selection. Ann N Y Acad Sci, 1239:25–32.

Zeidan MA, Igoe SA, Linnman C, Vitalo A, Levine JB, Klibanski A, Goldstein JM, Milad MR (2011) Estradiol modulates medial prefrontal cortex and amygdala activity during fear extinction in women and female rats. Biol Psychiatry, 70:920–927.

Zhang WN, Bast T, Feldon J (2001) The ventral hippocampus and fear conditioning in rats: different anterograde amnesias of fear after infusion of N-methyl-D-aspartate or its noncompetitive antagonist MK-801 into the ventral hippocampus. Behav Brain Res, 126:159–174.

Zhang WN, Bast T, Xu Y, Feldon J (2014) Temporary inhibition of dorsal or ventral hippocampus by muscimol: distinct effects on measures of innate anxiety on the elevated plus maze, but similar disruption of contextual fear conditioning. Behav Brain Res, 262:47–56.

Zhang Y, Xie B, Chen H, Li M, Guo X, Chen H (2016) Disrupted resting-state insular subregions functional connectivity in post-traumatic stress disorder. Medicine (Baltimore), 95:e4083.

Zimmermann KS, Li CC, Rainnie DG, Ressler KJ, Gourley SL (2018) Memory Retention Involves the Ventrolateral Orbitofrontal Cortex: Comparison with the Basolateral Amygdala. Neuropsychopharmacology, 43:674.